

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- (i) APPLICANT: GODDIJN, Oscar Johannes Maria  
PEN, Jan  
SMEEKENS, Josephus Christianus M.
- (ii) TITLE OF INVENTION: Regulating metabolism by modifying the  
level of trehalose-6-phosphate
- (iii) NUMBER OF SEQUENCES: 57
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: LADAS & PARRY
  - (B) STREET: 26 WEST 61 STREET
  - (C) CITY: NEW YORK
  - (D) STATE: NY
  - (E) COUNTRY: USA
  - (F) ZIP: 10023
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-Dos/MS-Dos
  - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER: --
  - (B) FILING DATE: 30-SEP-2003
  - (C) CLASSIFICATION: 435
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: 09/171,937
  - (B) FILING DATE: 28-APR-1999
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: PCT/EP97/02497
  - (B) FILING DATE: 02-MAY-1997
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: EP 96.201.225.8
  - (B) FILING DATE: 03-MAY-1996
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: EP 96.202.128.3
  - (B) FILING DATE: 26-JUL-1996
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: EP 96.202.395.8
  - (B) FILING DATE: 29-AUG-1996
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: MASS, Clifford J.

(A) TELEPHONE: (212) 708-1890  
(B) TELEFAX: (212) - 246-8959

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1450 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(ix) FEATURE:

(A) NAME/KEY: CDS  
(B) LOCATION: 21..1450

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```
ataaaaactct ccccggggacc atg act atg agt cgt tta gtc gta gta tct      50
                Met Thr Met Ser Arg Leu Val Val Val Ser
                  1                      5                      10

aac cgg att gca cca cca gac gag cac gcc gcc agt gcc ggt ggc ctt      98
Asn Arg Ile Ala Pro Pro Asp Glu His Ala Ala Ser Ala Gly Gly Leu
                  15                      20                      25

gcc gtt ggc ata ctg ggg gca ctg aaa gcc gca ggc gga ctg tgg ttt      146
Ala Val Gly Ile Leu Gly Ala Leu Lys Ala Ala Gly Gly Leu Trp Phe
                  30                      35                      40

ggc tgg agt ggt gaa aca ggg aat gag gat cag ccg cta aaa aag gtg      194
Gly Trp Ser Gly Glu Thr Gly Asn Glu Asp Gln Pro Leu Lys Lys Val
                  45                      50                      55

aaa aaa ggt aac att acg tgg gcc tct ttt aac ctc agc gaa cag gac      242
Lys Lys Gly Asn Ile Thr Trp Ala Ser Phe Asn Leu Ser Glu Gln Asp
                  60                      65                      70

ctt gac gaa tac tac aac caa ttc tcc aat gcc gtt ctc tgg ccc gct      290
Leu Asp Glu Tyr Tyr Asn Gln Phe Ser Asn Ala Val Leu Trp Pro Ala
                  75                      80                      85                      90

ttt cat tat cgg ctc gat ctg gtg caa ttt cag cgt cct gcc tgg gac      338
Phe His Tyr Arg Leu Asp Leu Val Gln Phe Gln Arg Pro Ala Trp Asp
                  95                      100                      105

ggc tat cta cgc gta aat gcg ttg ctg gca gat aaa tta ctg ccg ctg      386
Gly Tyr Leu Arg Val Asn Ala Leu Leu Ala Asp Lys Leu Leu Pro Leu
                  110                      115                      120

ttg caa gac gat gac att atc tgg atc cac gat tat cac ctg ttg cca      434
Leu Gln Asp Asp Asp Ile Ile Trp Ile His Asp Tyr His Leu Leu Pro
                  125                      130                      135

ttt gcg cat gaa tta cgc aaa cgg gga gtg aat aat cgc att ggt ttc      482
```

Phe	Ala	His	Glu	Leu	Arg	Lys	Arg	Gly	Val	Asn	Asn	Arg	Ile	Gly	Phe	
140						145				150						
ttt	ctg	cat	att	cct	ttc	ccg	aca	ccg	gaa	atc	ttc	aac	gcg	ctg	ccg	530
Phe	Leu	His	Ile	Pro	Phe	Pro	Thr	Pro	Glu	Ile	Phe	Asn	Ala	Leu	Pro	
155					160					165					170	
aca	tat	gac	acc	ttg	ctt	gaa	cag	ctt	tgt	gat	tat	gat	ttg	ctg	ggt	578
Thr	Tyr	Asp	Thr	Leu	Leu	Glu	Gln	Leu	Cys	Asp	Tyr	Asp	Leu	Leu	Gly	
				175					180					185		
ttc	cag	aca	gaa	aac	gat	cgt	ctg	gcg	ttc	ctg	gat	tgt	ctt	tct	aac	626
Phe	Gln	Thr	Glu	Asn	Asp	Arg	Leu	Ala	Phe	Leu	Asp	Cys	Leu	Ser	Asn	
			190					195					200			
ctg	acc	cgc	gtc	acg	aca	cgt	agc	gca	aaa	agc	cat	aca	gcc	tgg	ggc	674
Leu	Thr	Arg	Val	Thr	Thr	Arg	Ser	Ala	Lys	Ser	His	Thr	Ala	Trp	Gly	
		205					210					215				
aaa	gca	ttt	cga	aca	gaa	gtc	tac	ccg	atc	ggc	att	gaa	ccg	aaa	gaa	722
Lys	Ala	Phe	Arg	Thr	Glu	Val	Tyr	Pro	Ile	Gly	Ile	Glu	Pro	Lys	Glu	
	220					225					230					
ata	gcc	aaa	cag	gct	gcc	ggg	cca	ctg	ccg	cca	aaa	ctg	gcg	caa	ctt	770
Ile	Ala	Lys	Gln	Ala	Ala	Gly	Pro	Leu	Pro	Pro	Lys	Leu	Ala	Gln	Leu	
235					240					245					250	
aaa	gcg	gaa	ctg	aaa	aac	gta	caa	aat	atc	ttt	tct	gtc	gaa	cgg	ctg	818
Lys	Ala	Glu	Leu	Lys	Asn	Val	Gln	Asn	Ile	Phe	Ser	Val	Glu	Arg	Leu	
				255					260					265		
gat	tat	tcc	aaa	ggg	ttg	cca	gag	cgt	ttt	ctc	gcc	tat	gaa	gcg	ttg	866
Asp	Tyr	Ser	Lys	Gly	Leu	Pro	Glu	Arg	Phe	Leu	Ala	Tyr	Glu	Ala	Leu	
			270					275					280			
ctg	gaa	aaa	tat	ccg	cag	cat	cat	ggg	aaa	att	cgt	tat	acc	cag	att	914
Leu	Glu	Lys	Tyr	Pro	Gln	His	His	Gly	Lys	Ile	Arg	Tyr	Thr	Gln	Ile	
		285					290					295				
gca	cca	acg	tcg	cgt	ggg	gat	gtg	caa	gcc	tat	cag	gat	att	cgt	cat	962
Ala	Pro	Thr	Ser	Arg	Gly	Asp	Val	Gln	Ala	Tyr	Gln	Asp	Ile	Arg	His	
	300					305					310					
cag	ctc	gaa	aat	gaa	gct	gga	cga	att	aat	ggg	aaa	tac	ggg	caa	tta	1010
Gln	Leu	Glu	Asn	Glu	Ala	Gly	Arg	Ile	Asn	Gly	Lys	Tyr	Gly	Gln	Leu	
315					320					325					330	
ggc	tgg	acg	ccg	ctt	tat	tat	ttg	aat	cag	cat	ttt	gac	cgt	aaa	tta	1058
Gly	Trp	Thr	Pro	Leu	Tyr	Tyr	Leu	Asn	Gln	His	Phe	Asp	Arg	Lys	Leu	
				335					340					345		
ctg	atg	aaa	ata	ttc	cgc	tac	tct	gac	gtg	ggc	tta	gtg	acg	cca	ctg	1106
Leu	Met	Lys	Ile	Phe	Arg	Tyr	Ser	Asp	Val	Gly	Leu	Val	Thr	Pro	Leu	
			350					355					360			
cgt	gac	ggg	atg	aac	ctg	gta	gca	aaa	gag	tat	gtt	gct	gct	cag	gac	1154
Arg	Asp	Gly	Met	Asn	Leu	Val	Ala	Lys	Glu	Tyr	Val	Ala	Ala	Gln	Asp	

365

370

375

```

cca gcc aat ccg ggc gtt ctt gtt ctt tcg caa ttt gcg gga gcg gca 1202
Pro Ala Asn Pro Gly Val Leu Val Leu Ser Gln Phe Ala Gly Ala Ala
    380                      385                      390

aac gag tta acg tcg gcg tta att gtt aac ccc tac gat cgt gac gaa 1250
Asn Glu Leu Thr Ser Ala Leu Ile Val Asn Pro Tyr Asp Arg Asp Glu
    395                      400                      405                      410

gtt gca gct gcg ctg gat cgt gca ttg act atg tcg ctg gcg gaa cgt 1298
Val Ala Ala Ala Leu Asp Arg Ala Leu Thr Met Ser Leu Ala Glu Arg
                415                      420                      425

att tcc cgt cat gca gaa atg ctg gac gtt atc gtg aaa aac gat att 1346
Ile Ser Arg His Ala Glu Met Leu Asp Val Ile Val Lys Asn Asp Ile
                430                      435                      440

aac cac tgg cag gag tgc ttc att agc gac cta aag cag ata gtt ccg 1394
Asn His Trp Gln Glu Cys Phe Ile Ser Asp Leu Lys Gln Ile Val Pro
    445                      450                      455

cga agc gcg gaa agc cag cag cgc gat aaa gtt gct acc ttt cca aag 1442
Arg Ser Ala Glu Ser Gln Gln Arg Asp Lys Val Ala Thr Phe Pro Lys
    460                      465                      470

ctc tgc ag
Leu Cys
475

```

1450

## (2) INFORMATION FOR SEQ ID NO: 2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 476 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```

Met Thr Met Ser Arg Leu Val Val Val Ser Asn Arg Ile Ala Pro Pro
 1                      5                      10                      15

Asp Glu His Ala Ala Ser Ala Gly Gly Leu Ala Val Gly Ile Leu Gly
                20                      25                      30

Ala Leu Lys Ala Ala Gly Gly Leu Trp Phe Gly Trp Ser Gly Glu Thr
    35                      40                      45

Gly Asn Glu Asp Gln Pro Leu Lys Lys Val Lys Lys Gly Asn Ile Thr
    50                      55                      60

Trp Ala Ser Phe Asn Leu Ser Glu Gln Asp Leu Asp Glu Tyr Tyr Asn
    65                      70                      75                      80

```

Gln	Phe	Ser	Asn	Ala	Val	Leu	Trp	Pro	Ala	Phe	His	Tyr	Arg	Leu	Asp	85	90	95
Leu	Val	Gln	Phe	Gln	Arg	Pro	Ala	Trp	Asp	Gly	Tyr	Leu	Arg	Val	Asn	100	105	110
Ala	Leu	Leu	Ala	Asp	Lys	Leu	Leu	Pro	Leu	Leu	Gln	Asp	Asp	Asp	Ile	115	120	125
Ile	Trp	Ile	His	Asp	Tyr	His	Leu	Leu	Pro	Phe	Ala	His	Glu	Leu	Arg	130	135	140
Lys	Arg	Gly	Val	Asn	Asn	Arg	Ile	Gly	Phe	Phe	Leu	His	Ile	Pro	Phe	145	150	155
Pro	Thr	Pro	Glu	Ile	Phe	Asn	Ala	Leu	Pro	Thr	Tyr	Asp	Thr	Leu	Leu	165	170	175
Glu	Gln	Leu	Cys	Asp	Tyr	Asp	Leu	Leu	Gly	Phe	Gln	Thr	Glu	Asn	Asp	180	185	190
Arg	Leu	Ala	Phe	Leu	Asp	Cys	Leu	Ser	Asn	Leu	Thr	Arg	Val	Thr	Thr	195	200	205
Arg	Ser	Ala	Lys	Ser	His	Thr	Ala	Trp	Gly	Lys	Ala	Phe	Arg	Thr	Glu	210	215	220
Val	Tyr	Pro	Ile	Gly	Ile	Glu	Pro	Lys	Glu	Ile	Ala	Lys	Gln	Ala	Ala	225	230	235
Gly	Pro	Leu	Pro	Pro	Lys	Leu	Ala	Gln	Leu	Lys	Ala	Glu	Leu	Lys	Asn	245	250	255
Val	Gln	Asn	Ile	Phe	Ser	Val	Glu	Arg	Leu	Asp	Tyr	Ser	Lys	Gly	Leu	260	265	270
Pro	Glu	Arg	Phe	Leu	Ala	Tyr	Glu	Ala	Leu	Leu	Glu	Lys	Tyr	Pro	Gln	275	280	285
His	His	Gly	Lys	Ile	Arg	Tyr	Thr	Gln	Ile	Ala	Pro	Thr	Ser	Arg	Gly	290	295	300
Asp	Val	Gln	Ala	Tyr	Gln	Asp	Ile	Arg	His	Gln	Leu	Glu	Asn	Glu	Ala	305	310	315
Gly	Arg	Ile	Asn	Gly	Lys	Tyr	Gly	Gln	Leu	Gly	Trp	Thr	Pro	Leu	Tyr	325	330	335
Tyr	Leu	Asn	Gln	His	Phe	Asp	Arg	Lys	Leu	Leu	Met	Lys	Ile	Phe	Arg	340	345	350
Tyr	Ser	Asp	Val	Gly	Leu	Val	Thr	Pro	Leu	Arg	Asp	Gly	Met	Asn	Leu	355	360	365
Val	Ala	Lys	Glu	Tyr	Val	Ala	Ala	Gln	Asp	Pro	Ala	Asn	Pro	Gly	Val	370	375	380
Leu	Val	Leu	Ser	Gln	Phe	Ala	Gly	Ala	Ala	Asn	Glu	Leu	Thr	Ser	Ala			

385	390	395	400
Leu Ile Val Asn Pro Tyr Asp Arg Asp Glu Val Ala Ala Ala Leu Asp			
	405	410	415
Arg Ala Leu Thr Met Ser Leu Ala Glu Arg Ile Ser Arg His Ala Glu			
	420	425	430
Met Leu Asp Val Ile Val Lys Asn Asp Ile Asn His Trp Gln Glu Cys			
	435	440	445
Phe Ile Ser Asp Leu Lys Gln Ile Val Pro Arg Ser Ala Glu Ser Gln			
	450	455	460
Gln Arg Asp Lys Val Ala Thr Phe Pro Lys Leu Cys			
465	470	475	

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 835 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 18..818

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

ATAAACTCT CCCC GGG ATG ACA GAA CCG TTA ACC GAA ACC CCT GAA CTA	50
Met Thr Glu Pro Leu Thr Glu Thr Pro Glu Leu	
1 5 10	
TCC GCG AAA TAT GCC TGG TTT TTT GAT CTT GAT GGA ACG CTG GCG GAA	98
Ser Ala Lys Tyr Ala Trp Phe Phe Asp Leu Asp Gly Thr Leu Ala Glu	
15 20 25	
ATC AAA CCG CAT CCC GAT CAG GTC GTC GTG CCT GAC AAT ATT CTG CAA	146
Ile Lys Pro His Pro Asp Gln Val Val Val Pro Asp Asn Ile Leu Gln	
30 35 40	
GGA CTA CAG CTA CTG GCA ACC GCA AGT GAT GGT GCA TTG GCA TTG ATA	194
Gly Leu Gln Leu Leu Ala Thr Ala Ser Asp Gly Ala Leu Ala Leu Ile	
45 50 55	
TCA GGG CGC TCA ATG GTG GAG CTT GAC GCA CTG GCA AAA CCT TAT CGC	242
Ser Gly Arg Ser Met Val Glu Leu Asp Ala Leu Ala Lys Pro Tyr Arg	
60 65 70 75	
TTC CCG TTA GCG GGC GTG CAT GGG GCG GAG CGC CGT GAC ATC AAT GGT	290

Phe	Pro	Leu	Ala	Gly	Val	His	Gly	Ala	Glu	Arg	Arg	Asp	Ile	Asn	Gly		
				80					85					90			
AAA	ACA	CAT	ATC	GTT	CAT	CTG	CCG	GAT	GCG	ATT	GCG	CGT	GAT	ATT	AGC	338	
Lys	Thr	His	Ile	Val	His	Leu	Pro	Asp	Ala	Ile	Ala	Arg	Asp	Ile	Ser		
			95					100					105				
GTG	CAA	CTG	CAT	ACA	GTC	ATC	GCT	CAG	TAT	CCC	GGC	GCG	GAG	CTG	GAG	386	
Val	Gln	Leu	His	Thr	Val	Ile	Ala	Gln	Tyr	Pro	Gly	Ala	Glu	Leu	Glu		
		110					115					120					
GCG	AAA	GGG	ATG	GCT	TTT	GCG	CTG	CAT	TAT	CGT	CAG	GCT	CCG	CAG	CAT	434	
Ala	Lys	Gly	Met	Ala	Phe	Ala	Leu	His	Tyr	Arg	Gln	Ala	Pro	Gln	His		
	125					130				135							
GAA	GAC	GCA	TTA	ATG	ACA	TTA	GCG	CAA	CGT	ATT	ACT	CAG	ATC	TGG	CCA	482	
Glu	Asp	Ala	Leu	Met	Thr	Leu	Ala	Gln	Arg	Ile	Thr	Gln	Ile	Trp	Pro		
140					145				150					155			
CAA	ATG	GCG	TTA	CAG	CAG	GGA	AAG	TGT	GTT	GTC	GAG	ATC	AAA	CCG	AGA	530	
Gln	Met	Ala	Leu	Gln	Gln	Gly	Lys	Cys	Val	Val	Glu	Ile	Lys	Pro	Arg		
			160					165					170				
GGT	ACC	AGT	AAA	GGT	GAG	GCA	ATT	GCA	GCT	TTT	ATG	CAG	GAA	GCT	CCC	578	
Gly	Thr	Ser	Lys	Gly	Glu	Ala	Ile	Ala	Ala	Phe	Met	Gln	Glu	Ala	Pro		
			175					180					185				
TTT	ATC	GGG	CGA	ACG	CCC	GTA	TTT	CTG	GGC	GAT	GAT	TTA	ACC	GAT	GAA	626	
Phe	Ile	Gly	Arg	Thr	Pro	Val	Phe	Leu	Gly	Asp	Asp	Leu	Thr	Asp	Glu		
		190					195					200					
TCT	GGC	TTC	GCA	GTC	GTT	AAC	CGA	CTG	GGC	GGA	ATG	TCA	GTA	AAA	ATT	674	
Ser	Gly	Phe	Ala	Val	Val	Asn	Arg	Leu	Gly	Gly	Met	Ser	Val	Lys	Ile		
	205					210					215						
GGC	ACA	GGT	GCA	ACT	CAG	GCA	TCA	TGG	CGA	CTG	GCG	GGT	GTG	CCG	GAT	722	
Gly	Thr	Gly	Ala	Thr	Gln	Ala	Ser	Trp	Arg	Leu	Ala	Gly	Val	Pro	Asp		
220					225				230					235			
GTC	TGG	AGC	TGG	CTT	GAA	ATG	ATA	ACC	ACC	GCA	TTA	CAA	CAA	AAA	AGA	770	
Val	Trp	Ser	Trp	Leu	Glu	Met	Ile	Thr	Thr	Ala	Leu	Gln	Gln	Lys	Arg		
			240					245					250				
GAA	AAT	AAC	AGG	AGT	GAT	GAC	TAT	GAG	TCG	TTT	AGT	CGT	AGT	ATC	TAA	818	
Glu	Asn	Asn	Arg	Ser	Asp	Asp	Tyr	Glu	Ser	Phe	Ser	Arg	Ser	Ile			
		255					260					265					
CCGGATTGCA	CCTGCAG															835	

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 266 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Met Thr Glu Pro Leu Thr Glu Thr Pro Glu Leu Ser Ala Lys Tyr Ala  
1 5 10 15  
Trp Phe Phe Asp Leu Asp Gly Thr Leu Ala Glu Ile Lys Pro His Pro  
20 25 30  
Asp Gln Val Val Val Pro Asp Asn Ile Leu Gln Gly Leu Gln Leu Leu  
35 40 45  
Ala Thr Ala Ser Asp Gly Ala Leu Ala Leu Ile Ser Gly Arg Ser Met  
50 55 60  
Val Glu Leu Asp Ala Leu Ala Lys Pro Tyr Arg Phe Pro Leu Ala Gly  
65 70 75 80  
Val His Gly Ala Glu Arg Arg Asp Ile Asn Gly Lys Thr His Ile Val  
85 90 95  
His Leu Pro Asp Ala Ile Ala Arg Asp Ile Ser Val Gln Leu His Thr  
100 105 110  
Val Ile Ala Gln Tyr Pro Gly Ala Glu Leu Glu Ala Lys Gly Met Ala  
115 120 125  
Phe Ala Leu His Tyr Arg Gln Ala Pro Gln His Glu Asp Ala Leu Met  
130 135 140  
Thr Leu Ala Gln Arg Ile Thr Gln Ile Trp Pro Gln Met Ala Leu Gln  
145 150 155 160  
Gln Gly Lys Cys Val Val Glu Ile Lys Pro Arg Gly Thr Ser Lys Gly  
165 170 175  
Glu Ala Ile Ala Ala Phe Met Gln Glu Ala Pro Phe Ile Gly Arg Thr  
180 185 190  
Pro Val Phe Leu Gly Asp Asp Leu Thr Asp Glu Ser Gly Phe Ala Val  
195 200 205  
Val Asn Arg Leu Gly Gly Met Ser Val Lys Ile Gly Thr Gly Ala Thr  
210 215 220  
Gln Ala Ser Trp Arg Leu Ala Gly Val Pro Asp Val Trp Ser Trp Leu  
225 230 235 240  
Glu Met Ile Thr Thr Ala Leu Gln Gln Lys Arg Glu Asn Asn Arg Ser  
245 250 255  
Asp Asp Tyr Glu Ser Phe Ser Arg Ser Ile  
260 265

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 27 base pairs



- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

AAGCTTATGT TGCCATATAG AGTAGAT

27

(2) INFORMATION FOR SEQ ID NO: 6:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 29 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

GTAGTTGCCA TGGTGCAAAT GTTCATATG

29

(2) INFORMATION FOR SEQ ID NO: 7:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 26 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI--SENSE: NO

(ix) FEATURES:

- (A) NAME/KEY: modified base, Inosine
- (B) LOCATION: 4
- (C) OTHER INFORMATION: /note= N is Inosine

(ix) FEATURES:

- (A) NAME/KEY: modified base, Inosine
- (B) LOCATION: 6
- (C) OTHER INFORMATION: /note= N is Inosine

(ix) FEATURES:

- (A) NAME/KEY: modified base, Inosine
- (B) LOCATION: 9
- (C) OTHER INFORMATION: /note= N is Inosine

- (ix) FEATURES:
  - (A) NAME/KEY: modified base, Inosine
  - (B) LOCATION: 15
  - (C) OTHER INFORMATION: /note= N is Inosine

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

GAYNTNATAT GGRTNCAYGA YTAYCA

26

(2) INFORMATION FOR SEQ ID NO: 8:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 28 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (ix) FEATURES:
  - (A) NAME/KEY: modified base, Inosine
  - (B) LOCATION: 2
  - (C) OTHER INFORMATION: /note= N is Inosine
- (ix) FEATURES:
  - (A) NAME/KEY: modified base, Inosine
  - (B) LOCATION: 5
  - (C) OTHER INFORMATION: /note= N is Inosine
- (ix) FEATURES:
  - (A) NAME/KEY: modified base, Inosine
  - (B) LOCATION: 8
  - (C) OTHER INFORMATION: /note= N is Inosine
- (ix) FEATURES:
  - (A) NAME/KEY: modified base, Inosine
  - (B) LOCATION: 14
  - (C) OTHER INFORMATION: /note= N is Inosine
- (ix) FEATURES:
  - (A) NAME/KEY: modified base, Inosine
  - (B) LOCATION: 20
  - (C) OTHER INFORMATION: /note= N is Inosine
- (ix) FEATURES:
  - (A) NAME/KEY: modified base, Inosine
  - (B) LOCATION: 23
  - (C) OTHER INFORMATION: /note= N is Inosine

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

TNGGNTKNTT YYTNCAAYAYN CCNTTYCC

28

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURES:

- (A) NAME/KEY: modified base, Inosine
- (B) LOCATION: 3
- (C) OTHER INFORMATION: /note= N is Inosine

(ix) FEATURES:

- (A) NAME/KEY: modified base, Inosine
- (B) LOCATION: 6
- (C) OTHER INFORMATION: /note= N is Inosine

(ix) FEATURES:

- (A) NAME/KEY: modified base, Inosine
- (B) LOCATION: 15
- (C) OTHER INFORMATION: /note= N is Inosine

(ix) FEATURES:

- (A) NAME/KEY: modified base, Inosine
- (B) LOCATION: 21
- (C) OTHER INFORMATION: /note= N is Inosine

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

GYNACNARRT TCATNCCRTC NC

22

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 743 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Homo sapiens

(ix) FEATURE:

(A) NAME/KEY: CDS  
 (B) LOCATION: 1..743  
 (D) OTHER INFORMATION: /partial

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

GAC GTG ATG TGG ATG CAC GAC TAC CAT TTG ATG GTG TTG CCT ACG TTC	48
Asp Val Met Trp Met His Asp Tyr His Leu Met Val Leu Pro Thr Phe	
1 5 10 15	
TTG AGG AGG CGG TTC AAT CGT TTG AGA ATG GGG TTT TTC CTT CAC AGT	96
Leu Arg Arg Phe Asn Arg Leu Arg Met Gly Phe Phe Leu His Ser	
20 25 30	
CCA TTT CCC TCA TCT GAG ATT TAC AGG ACA CTT CCT GTT AGA GAG GAA	144
Pro Phe Pro Ser Ser Glu Ile Tyr Arg Thr Leu Pro Val Arg Glu Glu	
35 40 45	
ATA CTC AAG GCT TTG CTC TGT GCT GAC ATT GTT GGA TTC CAC ACT TTT	192
Ile Leu Lys Ala Leu Leu Cys Ala Asp Ile Val Gly Phe His Thr Phe	
50 55 60	
GAC TAC GCG AGA CAC TTC CTC TCT TGT TGC AGT CGG ATG TTG GGT TTA	240
Asp Tyr Ala Arg His Phe Leu Ser Cys Cys Ser Arg Met Leu Gly Leu	
65 70 75 80	
GAG TAT CAG TCT AAA AGA GGT TAT ATA GGG TTA GAA TAC TAT GGA CGG	288
Glu Tyr Gln Ser Lys Arg Gly Tyr Ile Gly Leu Glu Tyr Tyr Gly Arg	
85 90 95	
ACA GTA GGC ATC AAG ATT ATG CCC GTC GGG ATA CAT ATG GGT CAT ATT	336
Thr Val Gly Ile Lys Ile Met Pro Val Gly Ile His Met Gly His Ile	
100 105 110	
GAG TCC ATG AAG AAA CTT GCA GCG AAA GAG TTG ATG CTT AAG GCG CTA	384
Glu Ser Met Lys Lys Leu Ala Ala Lys Glu Leu Met Leu Lys Ala Leu	
115 120 125	
AAG CAG CAA TTT GAA GGG AAA ACT GTG TTG CTT GGT GCC GAT GAC CTG	432
Lys Gln Gln Phe Glu Gly Lys Thr Val Leu Leu Gly Ala Asp Asp Leu	
130 135 140	
GAT ATT TTC AAA GGT ATA AAC TTA AAG CTT CTA GCT ATG GAA CAG ATG	480
Asp Ile Phe Lys Gly Ile Asn Leu Lys Leu Leu Ala Met Glu Gln Met	
145 150 155 160	
CTC AAA CAG CAC CCC AAG TGG CAA GGG CAG GCT GTG TTG GTC CAG ATT	528
Leu Lys Gln His Pro Lys Trp Gln Gly Gln Ala Val Leu Val Gln Ile	
165 170 175	
GCA AAT CCT ACG AGG GGT AAA GGA GTA GAT TTT GAG GAA ATA CAG GCT	576
Ala Asn Pro Thr Arg Gly Lys Gly Val Asp Phe Glu Glu Ile Gln Ala	
180 185 190	
GAG ATA TCG GAA AGC TGT AAG AGA ATC AAT AAG CAA TTC GGC AAG CCT	624
Glu Ile Ser Glu Ser Cys Lys Arg Ile Asn Lys Gln Phe Gly Lys Pro	
195 200 205	

GGA TAT GAG CCT ATA GTT TAT ATT GAT AGG CCC GTG TCA AGC AGT GAA	672
Gly Tyr Glu Pro Ile Val Tyr Ile Asp Arg Pro Val Ser Ser Ser Glu	
210 215 220	
CGC ATG GCA TAT TAC AGT ATT GCA GAA TGT GTT GTT GTC ACG GCT GTG	720
Arg Met Ala Tyr Tyr Ser Ile Ala Glu Cys Val Val Val Thr Ala Val	
225 230 235 240	
AGC GAC GGC ATG AAC TTC GTC TC	743
Ser Asp Gly Met Asn Phe Val	
245	

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 247 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

Asp Val Met Trp Met His Asp Tyr His Leu Met Val Leu Pro Thr Phe	
1 5 10 15	
Leu Arg Arg Arg Phe Asn Arg Leu Arg Met Gly Phe Phe Leu His Ser	
20 25 30	
Pro Phe Pro Ser Ser Glu Ile Tyr Arg Thr Leu Pro Val Arg Glu Glu	
35 40 45	
Ile Leu Lys Ala Leu Leu Cys Ala Asp Ile Val Gly Phe His Thr Phe	
50 55 60	
Asp Tyr Ala Arg His Phe Leu Ser Cys Cys Ser Arg Met Leu Gly Leu	
65 70 75 80	
Glu Tyr Gln Ser Lys Arg Gly Tyr Ile Gly Leu Glu Tyr Tyr Gly Arg	
85 90 95	
Thr Val Gly Ile Lys Ile Met Pro Val Gly Ile His Met Gly His Ile	
100 105 110	
Glu Ser Met Lys Lys Leu Ala Ala Lys Glu Leu Met Leu Lys Ala Leu	
115 120 125	
Lys Gln Gln Phe Glu Gly Lys Thr Val Leu Leu Gly Ala Asp Asp Leu	
130 135 140	
Asp Ile Phe Lys Gly Ile Asn Leu Lys Leu Leu Ala Met Glu Gln Met	
145 150 155 160	
Leu Lys Gln His Pro Lys Trp Gln Gly Gln Ala Val Leu Val Gln Ile	
165 170 175	

Ala	Asn	Pro	Thr	Arg	Gly	Lys	Gly	Val	Asp	Phe	Glu	Glu	Ile	Gln	Ala	
			180					185						190		
Glu	Ile	Ser	Glu	Ser	Cys	Lys	Arg	Ile	Asn	Lys	Gln	Phe	Gly	Lys	Pro	
		195					200					205				
Gly	Tyr	Glu	Pro	Ile	Val	Tyr	Ile	Asp	Arg	Pro	Val	Ser	Ser	Ser	Glu	
	210					215					220					
Arg	Met	Ala	Tyr	Tyr	Ser	Ile	Ala	Glu	Cys	Val	Val	Val	Thr	Ala	Val	
225					230				235						240	
Ser	Asp	Gly	Met	Asn	Phe	Val										
					245											

(2) INFORMATION FOR SEQ ID NO: 12:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 395 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA to mRNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Nicotiana tabacum
  - (B) STRAIN: Samsun NN
  - (F) TISSUE TYPE: Leaf
- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION: 1..395
  - (D) OTHER INFORMATION: /partial
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

GCG	AAA	CCG	GTG	ATG	AAA	CTT	TAC	AGG	GAA	GCA	ACT	GAC	GGA	TCA	TAT	48
Ala	Lys	Pro	Val	Met	Lys	Leu	Tyr	Arg	Glu	Ala	Thr	Asp	Gly	Ser	Tyr	
1				5				10					15			
ATA	GAA	ACT	AAA	GAG	AGT	GCA	TTA	GTG	TGG	CAC	CAT	CAT	GAT	GCA	GAC	96
Ile	Glu	Thr	Lys	Glu	Ser	Ala	Leu	Val	Trp	His	His	His	Asp	Ala	Asp	
		20					25					30				
CCT	GAC	TTT	GGC	TCC	TGC	CAG	GCA	AAG	GAA	TTG	TTG	GAT	CAT	TTG	GAA	144
Pro	Asp	Phe	Gly	Ser	Cys	Gln	Ala	Lys	Glu	Leu	Leu	Asp	His	Leu	Glu	
		35				40						45				
AGC	GTA	CTT	GCA	AAT	GAA	CCT	GCA	GTT	GTT	AAG	AGG	GGC	CAA	CAT	ATT	192
Ser	Val	Leu	Ala	Asn	Glu	Pro	Ala	Val	Val	Lys	Arg	Gly	Gln	His	Ile	
	50					55				60						

GTT GAA GTC AAG CCA CAA GGT GTG ACC AAA GGA TTA GTT TCA GAG AAG	240
Val Glu Val Lys Pro Gln Gly Val Thr Lys Gly Leu Val Ser Glu Lys	
65 70 75 80	
GTT CTC TCG ATG ATG GTT GAT AGT GGG AAA CCG CCC GAT TTT GTT ATG	288
Val Leu Ser Met Met Val Asp Ser Gly Lys Pro Pro Asp Phe Val Met	
85 90 95	
TGC ATT GGA GAT GAT AGG TCA GAC GAA GAC ATG TTT GAG AGC ATA TTA	336
Cys Ile Gly Asp Asp Arg Ser Asp Glu Asp Met Phe Glu Ser Ile Leu	
100 105 110	
AGC ACC GTA TCC AGT CTG TCA GTC ACT GCT GCC CCT GAT GTC TTT GCC	384
Ser Thr Val Ser Ser Leu Ser Val Thr Ala Ala Pro Asp Val Phe Ala	
115 120 125	
TGC ACC GTC GG	395
Cys Thr Val	
130	

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 131 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

Ala Lys Pro Val Met Lys Leu Tyr Arg Glu Ala Thr Asp Gly Ser Tyr	
1 5 10 15	
Ile Glu Thr Lys Glu Ser Ala Leu Val Trp His His His Asp Ala Asp	
20 25 30	
Pro Asp Phe Gly Ser Cys Gln Ala Lys Glu Leu Leu Asp His Leu Glu	
35 40 45	
Ser Val Leu Ala Asn Glu Pro Ala Val Val Lys Arg Gly Gln His Ile	
50 55 60	
Val Glu Val Lys Pro Gln Gly Val Thr Lys Gly Leu Val Ser Glu Lys	
65 70 75 80	
Val Leu Ser Met Met Val Asp Ser Gly Lys Pro Pro Asp Phe Val Met	
85 90 95	
Cys Ile Gly Asp Asp Arg Ser Asp Glu Asp Met Phe Glu Ser Ile Leu	
100 105 110	
Ser Thr Val Ser Ser Leu Ser Val Thr Ala Ala Pro Asp Val Phe Ala	
115 120 125	
Cys Thr Val	
130	

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 491 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Nicotiana tabacum
- (B) STRAIN: Samsun NN
- (F) TISSUE TYPE: Leaf

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..491
- (D) OTHER INFORMATION: /partial

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

GGG	CTG	TCG	GCG	GAA	CAC	GGC	TAT	TTC	TTG	AGG	ACG	AGT	CAA	GAT	GAA	48
Gly	Leu	Ser	Ala	Glu	His	Gly	Tyr	Phe	Leu	Arg	Thr	Ser	Gln	Asp	Glu	
1				5					10					15		
GAA	TGG	GAA	ACA	TGT	GTA	CCA	CCA	GTG	GAA	TGT	TGT	TGG	AAA	GAA	ATA	96
Glu	Trp	Glu	Thr	Cys	Val	Pro	Pro	Val	Glu	Cys	Cys	Trp	Lys	Glu	Ile	
			20					25					30			



GCT GAG CCT GTT ATG CAA CTT TAC ACT GAG ACT ACT GAT GGA TCA GTT Ala Glu Pro Val Met Gln Leu Tyr Thr Glu Thr Thr Asp Gly Ser Val	144
35 40 45	
ATT GAA GAT AAG GAA ACA TCA ATG GTC TGG TCT TAC GAG GAT GCG GAT Ile Glu Asp Lys Glu Thr Ser Met Val Trp Ser Tyr Glu Asp Ala Asp	192
50 55 60	
CCT GAT TTT GGA TCA TGT CAG GCT AAG GAA CTT CTT GAT CAC CTA GAA Pro Asp Phe Gly Ser Cys Gln Ala Lys Glu Leu Leu Asp His Leu Glu	240
65 70 75 80	
AGT GTA CTA GCT AAT GAA CCG GTC ACT GTC AGG AGT GGA CAG AAT ATA Ser Val Leu Ala Asn Glu Pro Val Thr Val Arg Ser Gly Gln Asn Ile	288
85 90 95	
GTG GAA GTT AAG CCC CAG GGT GTA TCC AAA GGG CTT GTT GCC AAG CGC Val Glu Val Lys Pro Gln Gly Val Ser Lys Gly Leu Val Ala Lys Arg	336
100 105 110	
CTG CTT TCC GCA ATG CAA GAG AAA GGA ATG TCA CCA GAT TTT GTC CTT Leu Leu Ser Ala Met Gln Glu Lys Gly Met Ser Pro Asp Phe Val Leu	384
115 120 125	
TGC ATA GGA GAT GAC CGA TCG GAT GAA GAC ATG TTC GAG GTG ATC ATG Cys Ile Gly Asp Asp Arg Ser Asp Glu Asp Met Phe Glu Val Ile Met	432
130 135 140	
AGC TCG ATG TCT GGC CCG TCC ATG GCT CCA ACA GCT GAA GTC TTT GCC Ser Ser Met Ser Gly Pro Ser Met Ala Pro Thr Ala Glu Val Phe Ala	480
145 150 155 160	
TGC ACC GTC GG Cys Thr Val	491

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 163 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

Gly Leu Ser Ala Glu His Gly Tyr Phe Leu Arg Thr Ser Gln Asp Glu 1 5 10 15
Glu Trp Glu Thr Cys Val Pro Pro Val Glu Cys Cys Trp Lys Glu Ile 20 25 30
Ala Glu Pro Val Met Gln Leu Tyr Thr Glu Thr Thr Asp Gly Ser Val 35 40 45
Ile Glu Asp Lys Glu Thr Ser Met Val Trp Ser Tyr Glu Asp Ala Asp 50 55 60

Pro	Asp	Phe	Gly	Ser	Cys	Gln	Ala	Lys	Glu	Leu	Leu	Asp	His	Leu	Glu
65					70					75					80
Ser	Val	Leu	Ala	Asn	Glu	Pro	Val	Thr	Val	Arg	Ser	Gly	Gln	Asn	Ile
				85					90					95	
Val	Glu	Val	Lys	Pro	Gln	Gly	Val	Ser	Lys	Gly	Leu	Val	Ala	Lys	Arg
			100					105					110		
Leu	Leu	Ser	Ala	Met	Gln	Glu	Lys	Gly	Met	Ser	Pro	Asp	Phe	Val	Leu
		115						120				125			
Cys	Ile	Gly	Asp	Asp	Arg	Ser	Asp	Glu	Asp	Met	Phe	Glu	Val	Ile	Met
	130						135				140				
Ser	Ser	Met	Ser	Gly	Pro	Ser	Met	Ala	Pro	Thr	Ala	Glu	Val	Phe	Ala
145					150					155					160
Cys	Thr	Val													

(2) INFORMATION FOR SEQ ID NO: 16:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 361 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA to mRNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Nicotiana tabacum*
  - (B) STRAIN: Samsun NN
  - (F) TISSUE TYPE: Leaf

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

TTTGATTATG ATGGGACGCT GCTGTCTGGAG GAGAGTGTGG ACAAACCCC GAGTGAAGAT	60
GACATCTCAA TTCTGAATGG TTTATGCAGT GATCCAAAGA ACGTAGTCTT TATCGTGAGT	120
GGCAGAGGAA AGGATACTACT TAGCAAGTGG TTCTCTCCGT GTCCGAGACT CGGCCTATCA	180
GCAGAACATG GATATTTTAC TAGGTGGAGT AAGGATTCCG AGTGGGAATC TCGTCCATAG	240
CTGCAGACCT TGACTGGAAA AAAATAGTGT TGCCTATTAT GGAGCGCTAC ACAGAGCACA	300
GATGGTTCGT CGATAGAACA GAAGGAAACC TCGTGTGGC TCATCAAATG CTGGCCCCGA	360

(2) INFORMATION FOR SEQ ID NO: 17:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 118 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA to mRNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Nicotiana tabacum
  - (B) STRAIN: Samsun NN
  - (F) TISSUE TYPE: Leaf

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

```
GGAAACCCAC AGGATGTAAG CAAAGTTTTA GTTTTTGAGA TCTCTTGGCA TCAAGCAAAG    60
TAGAGGGAAG TCACCCGATT CGTGCTGTGC GTAGGGATGA CAGATCGGAC GACTTAGA    118
```

(2) INFORMATION FOR SEQ ID NO: 18:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 417 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA to mRNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Nicotiana tabacum
  - (B) STRAIN: Samsun NN
  - (F) TISSUE TYPE: Leaf

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

```
TTGTGGCCGA TGTCCACTA CATGTTGCCG TTCTCACCTG ACCATGGAGG CCGCTTTGAT    60
CGCTCTATGT GGGAAGCATA TGTTTCTGCC AACAAGTTGT TTTCACAAAA AGTAGTTGAG    120
GTTCTTAATC CTGAGGATGA CTTTGTCTGG ATTCATGATT ATCATTTGAT GGTGTTGCCA    180
ACGTTCTTGA GGAGGCGGTT CAATCGTTTG AGAATGGGGT TTTTCCTTCA CAGTCCATTC    240
CTTCATCTGA GATTACAGG ACACTTCCTG TTAGAGAGGA AATACTCAAG GCTTTGCTCT    300
GTGCTGACAT TGTTGGATTC CACACTTTTG ACTACGCGAG ACACTTCCTC TCTTGTTGCA    360
```

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 411 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Nicotiana tabacum*
- (B) STRAIN: Samsun NN
- (F) TISSUE TYPE: Leaf

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

```
GGGTCATATT GATCCATGAA GAAATTGCAG CGAAAGAGTG ATGCTTTAAT GCGTAAAGCA    60
GCAATTTGAA GGGAAAACCTG TGTTGTTAGG TGCCGATGAC CTGGATATTT TCAAAGGTAT   120
GAACTTAAAG CTTCTAGCTA TGGAACAGAT GCTCAAACAT CACCCCAAGT GGCAAGGGCA   180
GGCTGTGTTG GTCCAAGATT GCAAATCCTA CGAGGGGTAA AGGAGTAGAT TTTGACGAAA   240
TACGGCTGAG ACATCGGAAA GCTGTAAGAG AATCAATAAG CAATTCGGCA AGCCTGGATA   300
TGAGCCTATA GTTTATATTG ATAGGCCCGT GTCAAGCAGT GAACGCATGG CATATTACAG   360
TATTGCAGGA TGTGTTGTGG TCACGCTGTG AGCGATGGCA TGAATCTGTT C             411
```

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 405 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Nicotiana tabacum*
- (B) STRAIN: Samsun NN
- (F) TISSUE TYPE: Leaf

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

TGGGGTGGTT CCTGCATACG CCGTTTCCTT CTTCTGAGAT ATATAAACT TTGCCTATTC 60  
GCGAAAGATC TTACAGCTCT CTTGAATTCA ATTTGATTGG GTTCCACACT TTTGACTATG 120  
CAGGCACTTC CTCTCGTGTT GCAGTCGGAT GTTAGGTATT TCTTATGATC AAAAAGGGGT 180  
TACATAGGCC TCGATATTAT GGCAGGACTG TAATATAAAA ATTCTGCCAG CGGGTATTCA 240  
TATGGGGCAG CTTCAAGCAAG TCTTGAGTCT TCCTGAAACG GAGGCAAAAT CTCGGAAGTC 300  
GTGCAGCATT TAATCATCAG GGGGAGGACA TTGTTGCTGG GATTGATGAC TGGACATATT 360  
TAAAGGCTCA TTTGAATTTA TTACCATGGA ACAACTCTAT TGCAC 405

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 427 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Nicotiana tabacum*
- (B) STRAIN: Samsun NN
- (F) TISSUE TYPE: Leaf

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

ATCATATGGG GCAGCTTCAG CAATCTTGAT CTTCTGAAA CGGAGGCAAA AGTCTTCGGA 60  
ACTCGGCAGC AGTTTAATCA TCAGGGGAGG ACATTGTTGC TGGGAGTTGA TGACATGGAC 120  
ATATTTAAAG GCATCAGTTT GAAGTTATTA GCAATGGAAC AACTTCTATT GCAGCACCCG 180  
GAGAAGCAGG GGAAGGTTGT TTTGGTGCAG ATAGCCAATC CTGCTAGAGG CAAAGGAAAA 240  
GATGTCAAAG AAGTGCAGGA AGAACTCAT TGACGGTGAA GCGAATTAAT GAAGCATTTG 300  
GAAGACCTGG GTACGAACCA GTTATCTTGA TTGATAAGCC ACTAAAGTTT TATGAAAGGA 360  
TTGCTTATTA TGTGTTTGCA GAGTGTGACC TAGTCACTGC TGTCAGCGAT GGCATGAACC 420  
TCGTCTC 427

(2) INFORMATION FOR SEQ ID NO: 22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 315 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Nicotiana tabacum
- (B) STRAIN: Samsun NN
- (F) TISSUE TYPE: Leaf

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

```
GATGTGGATG CATGACTACC AATCCAAGAG GGGGTATATT GGTCTTGACT ATTATGGTAA   60
ACTGTGACCA TTAAATCCT TCCAGTTGGT ATTCACATGG GACAACTCCA AAATGTTATG  120
TCACTACAGA CACGGGAAAG AAAGCAAAGG AGTTGAAAGA AAAATATGAG GGGAAAATTG  180
TGATGTTAGG TATTGATGAT ATGGACATGT TTAAAGGAAT TGGTCTAAAG TTTCTGGCAA  240
TGGGGAGGCT TCTAGATGAA AACCTGTCT TGAGGGGTAA AGTGGTATTG GTTCAATCAC  300
CAGGCCTGGA AATTA                                                    315
```

(2) INFORMATION FOR SEQ ID NO: 23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 352 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Nicotiana tabacum
- (B) STRAIN: Samsun NN
- (F) TISSUE TYPE: Leaf

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

```
AGAAGTAAAG GGAGTGAGTC CCCGAGGTTT AAAAAGAGGT CAACAGAATT GCAGTGAAAT   60
TAATAAAAAA TATGGCAAAC CGGGGTACAA GCCGATTGTT TGTATCAATG GTCCAGTTTC  120
GACACAAGAC AAGATTGCAC ATTATGCGGT CTTGAGTGTG TTGTTGTAA TGCTGTTAGA  180
```



TGG GAC CAA TAT TAT CAC TGC TTT TGT AAG CAG TAT TTG TGG CCG ATA Trp Asp Gln Tyr Tyr His Cys Phe Cys Lys Gln Tyr Leu Trp Pro Ile	368
55 60 65	
TTT CAT TAC AAG GTT CCC GCT TCT GAC GTC AAG AGT GTC CCG AAT AGT Phe His Tyr Lys Val Pro Ala Ser Asp Val Lys Ser Val Pro Asn Ser	416
70 75 80	
CGG GAT TCA TGG AAC GCT TAT GTT CAC GTG AAC AAA GAG TTT TCC CAG Arg Asp Ser Trp Asn Ala Tyr Val His Val Asn Lys Glu Phe Ser Gln	464
85 90 95	
AAG GTG ATG GAG GCA GTA ACC AAT GCT AGC AAT TAT GTA TGG ATA CAT Lys Val Met Glu Ala Val Thr Asn Ala Ser Asn Tyr Val Trp Ile His	512
100 105 110	
GAC TAC CAT TTA ATG ACG CTA CCG ACT TTC TTG AGG CCG GAT TTT TGT Asp Tyr His Leu Met Thr Leu Pro Thr Phe Leu Arg Arg Asp Phe Cys	560
115 120 125 130	
CGT TTT AAA ATC GGT TTT TTT CTG CAT AGC CCG TTT CCT TCC TCG GAG Arg Phe Lys Ile Gly Phe Phe Leu His Ser Pro Phe Pro Ser Ser Glu	608
135 140 145	
GTT TAC AAG ACC CTA CCA ATG AGA AAC GAG CTC TTG AAG GGT CTG TTA Val Tyr Lys Thr Leu Pro Met Arg Asn Glu Leu Leu Lys Gly Leu Leu	656
150 155 160	
AAT GCT GAT CTT ATC GGG TTC CAT ACA TAC GAT TAT GCC CGT CAT TTT Asn Ala Asp Leu Ile Gly Phe His Thr Tyr Asp Tyr Ala Arg His Phe	704
165 170 175	
CTA ACG TGT TGT AGT CGA ATG TTT GGT TTG GAT CAT CAG TTG AAA AGG Leu Thr Cys Cys Ser Arg Met Phe Gly Leu Asp His Gln Leu Lys Arg	752
180 185 190	
GGG TAC ATT TTC TTG GAA TAT AAT GGA AGG AGC ATT GAG ATC AAG ATA Gly Tyr Ile Phe Leu Glu Tyr Asn Gly Arg Ser Ile Glu Ile Lys Ile	800
195 200 205 210	
AAG GCG AGC GGG ATT CAT GTT GGT CGA ATG GAG TCG TAC TTG AGT CAG Lys Ala Ser Gly Ile His Val Gly Arg Met Glu Ser Tyr Leu Ser Gln	848
215 220 225	
CCC GAT ACA AGA TTA CAA GTT CAA GAA CTA AAA AAA CGT TTC GAA GGG Pro Asp Thr Arg Leu Gln Val Gln Glu Leu Lys Lys Arg Phe Glu Gly	896
230 235 240	
AAA ATC GTG CTA CTT GGA GTT GAT GAT TTG GAT ATA TTC AAA GGT GTG Lys Ile Val Leu Leu Gly Val Asp Asp Leu Asp Ile Phe Lys Gly Val	944
245 250 255	
AAC TTC AAG GTT TTA GCG TTG GAG AAG TTA CTT AAA TCA CAC CCG AGT Asn Phe Lys Val Leu Ala Leu Glu Lys Leu Leu Lys Ser His Pro Ser	992
260 265 270	



TGG CAA GGG CGT GTG GTT TTG GTG CAA ATC TTG AAT CCC GCT CGC GCG Trp Gln Gly Arg Val Val Leu Val Gln Ile Leu Asn Pro Ala Arg Ala 275 280 285 290	1040
CGT TGC CAA GAC GTC GAT GAG ATC AAT GCC GAG ATA AGA ACA GTC TGT Arg Cys Gln Asp Val Asp Glu Ile Asn Ala Glu Ile Arg Thr Val Cys 295 300 305	1088
GAA AGA ATC AAT AAC GAA CTG GGA AGC CCG GGA TAC CAG CCC GTT GTG Glu Arg Ile Asn Asn Glu Leu Gly Ser Pro Gly Tyr Gln Pro Val Val 310 315 320	1136
TTA ATT GAT GGG CCC GTT TCG TTA AGT GAA AAA GCT GCT TAT TAT GCT Leu Ile Asp Gly Pro Val Ser Leu Ser Glu Lys Ala Ala Tyr Tyr Ala 325 330 335	1184
ATC GCC GAT ATG GCA ATT GTT ACA CCG TTA CGT GAC GGC ATG AAT CTT Ile Ala Asp Met Ala Ile Val Thr Pro Leu Arg Asp Gly Met Asn Leu 340 345 350	1232
ATC CCG TAC GAG TAC GTC GTT TCC CGA CAA AGT GTT AAT GAC CCA AAT Ile Pro Tyr Glu Tyr Val Val Ser Arg Gln Ser Val Asn Asp Pro Asn 355 360 365 370	1280
CCC AAT ACT CCA AAA AAG AGC ATG CTA GTG GTC TCC GAG TTC ATC GGG Pro Asn Thr Pro Lys Lys Ser Met Leu Val Val Ser Glu Phe Ile Gly 375 380 385	1328
TGT TCA CTA TCT TTA ACC GGG GCC ATA CGG GTC AAC CCA TGG GAT GAG Cys Ser Leu Ser Leu Thr Gly Ala Ile Arg Val Asn Pro Trp Asp Glu 390 395 400	1376
TTG GAG ACA GCA GAA GCA TTA TAC GAC GCA CTC ATG GCT CCT GAT GAC Leu Glu Thr Ala Glu Ala Leu Tyr Asp Ala Leu Met Ala Pro Asp Asp 405 410 415	1424
CAT AAA GAA ACC GCC CAC ATG AAA CAG TAT CAA TAC ATT ATC TCC CAT His Lys Glu Thr Ala His Met Lys Gln Tyr Gln Tyr Ile Ile Ser His 420 425 430	1472
GAT GTA GCT AAC TGG GCT CGT AGC TTC TTT CAA GAT TTA GAG CAA GCG Asp Val Ala Asn Trp Ala Arg Ser Phe Phe Gln Asp Leu Glu Gln Ala 435 440 445 450	1520
TGC ATC GAT CAT TCT CGT AAA CGA TGC ATG AAT TTA GGA TTT GGG TTA Cys Ile Asp His Ser Arg Lys Arg Cys Met Asn Leu Gly Phe Gly Leu 455 460 465	1568
GAT ACT AGA GTC GTT CTT TTT GAT GAG AAG TTT AGC AAG TTG GAT ATA Asp Thr Arg Val Val Leu Phe Asp Glu Lys Phe Ser Lys Leu Asp Ile 470 475 480	1616
GAT GTC TTG GAG AAT GCT TAT TCC ATG GCT CAA AAT CGG GCC ATA CTT Asp Val Leu Glu Asn Ala Tyr Ser Met Ala Gln Asn Arg Ala Ile Leu 485 490 495	1664

TTG GAC TAT GAC GGC ACT GTT ACT CCA TCT ATC AGT AAA TCT CCA ACT Leu Asp Tyr Asp Gly Thr Val Thr Pro Ser Ile Ser Lys Ser Pro Thr 500 505 510	1712
GAA GCT GTT ATC TCC ATG ATC AAC AAA CTG TGC AAT GAT CCA AAG AAC Glu Ala Val Ile Ser Met Ile Asn Lys Leu Cys Asn Asp Pro Lys Asn 515 520 525 530	1760
ATG GTG TTC ATC GTT AGT GGA CGC AGT AGA GAA AAT CTT GGC AGT TGG Met Val Phe Ile Val Ser Gly Arg Ser Arg Glu Asn Leu Gly Ser Trp 535 540 545	1808
TTC GGC GCG TGT GAG AAA CCC GCC ATT GCA GCT GAG CAC GGA TAC TTT Phe Gly Ala Cys Glu Lys Pro Ala Ile Ala Ala Glu His Gly Tyr Phe 550 555 560	1856
ATA AGG TGG GCG GGT GAT CAA GAA TGG GAA ACG TGC GCA CGT GAG AAT Ile Arg Trp Ala Gly Asp Gln Glu Trp Glu Thr Cys Ala Arg Glu Asn 565 570 575	1904
AAT GTC GGG TGG ATG GAA ATG GCT GAG CCG GTT ATG AAT CTT TAT ACA Asn Val Gly Trp Met Glu Met Ala Glu Pro Val Met Asn Leu Tyr Thr 580 585 590	1952
GAA ACT ACT GAC GGT TCG TAT ATT GAA AAG AAA GAA ACT GCA ATG GTT Glu Thr Thr Asp Gly Ser Tyr Ile Glu Lys Lys Glu Thr Ala Met Val 595 600 605 610	2000
TGG CAC TAT GAA GAT GCT GAT AAA GAT CTT GGG TTG GAG CAG GCT AAG Trp His Tyr Glu Asp Ala Asp Lys Asp Leu Gly Leu Glu Gln Ala Lys 615 620 625	2048
GAA CTG TTG GAC CAT CTT GAA AAC GTG CTC GCT AAT GAG CCC GTT GAA Glu Leu Leu Asp His Leu Glu Asn Val Leu Ala Asn Glu Pro Val Glu 630 635 640	2096
GTG AAA CGA GGT CAA TAC ATT GTA GAA GTT AAA CCA CAG GTA CCC CAT Val Lys Arg Gly Gln Tyr Ile Val Glu Val Lys Pro Gln Val Pro His 645 650 655	2144
GGG TTA CCT TCT TGT TAT GAC ATT CAT AGG CAC AGA TTT GTA GAA TCT Gly Leu Pro Ser Cys Tyr Asp Ile His Arg His Arg Phe Val Glu Ser 660 665 670	2192
TTT AAC TTA AAT TTC TTT AAA TAT GAA TGC AAT TAT AGG GGG TCA CTG Phe Asn Leu Asn Phe Phe Lys Tyr Glu Cys Asn Tyr Arg Gly Ser Leu 675 680 685 690	2240
AAA GGT ATA GTT GCA GAG AAG ATT TTT GCG TTC ATG GCT GAA AAG GGA Lys Gly Ile Val Ala Glu Lys Ile Phe Ala Phe Met Ala Glu Lys Gly 695 700 705	2288
AAA CAG GCT GAT TTC GTG TTG AGC GTT GGA GAT GAT AGA AGT GAT GAA Lys Gln Ala Asp Phe Val Leu Ser Val Gly Asp Asp Arg Ser Asp Glu 710 715 720	2336

GAC ATG TTT GTG GCC ATT GGG GAT GGA ATA AAA AAG GGT CGG ATA ACT 2384  
 Asp Met Phe Val Ala Ile Gly Asp Gly Ile Lys Lys Gly Arg Ile Thr  
 725 730 735

AAC AAC AAT TCA GTG TTT ACA TGC GTA GTG GGA GAG AAA CCG AGT GCA 2432  
 Asn Asn Asn Ser Val Phe Thr Cys Val Val Gly Glu Lys Pro Ser Ala  
 740 745 750

GCT GAG TAC TTT TTA GAC GAG ACG AAA GAT GTT TCA ATG ATG CTC GAG 2480  
 Ala Glu Tyr Phe Leu Asp Glu Thr Lys Asp Val Ser Met Met Leu Glu  
 755 760 765 770

AAG CTC GGG TGT CTC AGC AAC CAA GGA T GATGATCCGG AAGCTTCTCG 2528  
 Lys Leu Gly Cys Leu Ser Asn Gln Gly  
 775

TGATCTTTAT GAGTTAAAAAG TTTTCGACTT TTTCTTCATC AAGATTCATG GGAAAGTTGT 2588

TCAATATGAA CTTGTGTTTC TTGGTTCTGG ATTTTAGGGA GTCTATGGAT CC 2640

(2) INFORMATION FOR SEQ ID NO: 25:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 779 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

Met His Ile Lys Asp Ala Leu Pro Ala Ala Val Glu Val Phe Tyr Val  
 1 5 10 15

Gly Ala Leu Arg Ala Asp Val Gly Pro Thr Glu Gln Asp Asp Val Ser  
 20 25 30

Lys Thr Leu Leu Asp Arg Phe Asn Cys Val Ala Val Phe Val Pro Thr  
 35 40 45

Ser Lys Trp Asp Gln Tyr Tyr His Cys Phe Cys Lys Gln Tyr Leu Trp  
 50 55 60

Pro Ile Phe His Tyr Lys Val Pro Ala Ser Asp Val Lys Ser Val Pro  
 65 70 75 80

Asn Ser Arg Asp Ser Trp Asn Ala Tyr Val His Val Asn Lys Glu Phe  
 85 90 95

Ser Gln Lys Val Met Glu Ala Val Thr Asn Ala Ser Asn Tyr Val Trp  
 100 105 110

Ile His Asp Tyr His Leu Met Thr Leu Pro Thr Phe Leu Arg Arg Asp  
 115 120 125

Phe Cys Arg Phe Lys Ile Gly Phe Phe Leu His Ser Pro Phe Pro Ser  
 130 135 140

Ser Glu Val Tyr Lys Thr Leu Pro Met Arg Asn Glu Leu Leu Lys Gly  
 145 150 155 160

Leu Leu Asn Ala Asp Leu Ile Gly Phe His Thr Tyr Asp Tyr Ala Arg  
 165 170 175

His Phe Leu Thr Cys Cys Ser Arg Met Phe Gly Leu Asp His Gln Leu  
 180 185 190

Lys Arg Gly Tyr Ile Phe Leu Glu Tyr Asn Gly Arg Ser Ile Glu Ile  
 195 200 205

Lys Ile Lys Ala Ser Gly Ile His Val Gly Arg Met Glu Ser Tyr Leu  
 210 215 220

Ser Gln Pro Asp Thr Arg Leu Gln Val Gln Glu Leu Lys Lys Arg Phe  
 225 230 235 240

Glu Gly Lys Ile Val Leu Leu Gly Val Asp Asp Leu Asp Ile Phe Lys  
 245 250 255

Gly Val Asn Phe Lys Val Leu Ala Leu Glu Lys Leu Leu Lys Ser His  
 260 265 270

Pro Ser Trp Gln Gly Arg Val Val Leu Val Gln Ile Leu Asn Pro Ala  
 275 280 285

Arg Ala Arg Cys Gln Asp Val Asp Glu Ile Asn Ala Glu Ile Arg Thr  
 290 295 300

Val Cys Glu Arg Ile Asn Asn Glu Leu Gly Ser Pro Gly Tyr Gln Pro  
 305 310 315 320

Val Val Leu Ile Asp Gly Pro Val Ser Leu Ser Glu Lys Ala Ala Tyr  
 325 330 335

Tyr Ala Ile Ala Asp Met Ala Ile Val Thr Pro Leu Arg Asp Gly Met  
 340 345 350

Asn Leu Ile Pro Tyr Glu Tyr Val Val Ser Arg Gln Ser Val Asn Asp  
 355 360 365

Pro Asn Pro Asn Thr Pro Lys Lys Ser Met Leu Val Val Ser Glu Phe  
 370 375 380

Ile Gly Cys Ser Leu Ser Leu Thr Gly Ala Ile Arg Val Asn Pro Trp  
 385 390 395 400

Asp Glu Leu Glu Thr Ala Glu Ala Leu Tyr Asp Ala Leu Met Ala Pro  
 405 410 415

Asp Asp His Lys Glu Thr Ala His Met Lys Gln Tyr Gln Tyr Ile Ile  
 420 425 430

Ser His Asp Val Ala Asn Trp Ala Arg Ser Phe Phe Gln Asp Leu Glu  
 435 440 445

Gln Ala Cys Ile Asp His Ser Arg Lys Arg Cys Met Asn Leu Gly Phe  
 450 455 460

Gly Leu Asp Thr Arg Val Val Leu Phe Asp Glu Lys Phe Ser Lys Leu  
 465 470 475 480

Asp Ile Asp Val Leu Glu Asn Ala Tyr Ser Met Ala Gln Asn Arg Ala  
 485 490 495

Ile Leu Leu Asp Tyr Asp Gly Thr Val Thr Pro Ser Ile Ser Lys Ser  
 500 505 510

Pro Thr Glu Ala Val Ile Ser Met Ile Asn Lys Leu Cys Asn Asp Pro  
 515 520 525

Lys Asn Met Val Phe Ile Val Ser Gly Arg Ser Arg Glu Asn Leu Gly  
 530 535 540

Ser Trp Phe Gly Ala Cys Glu Lys Pro Ala Ile Ala Ala Glu His Gly  
 545 550 555 560

Tyr Phe Ile Arg Trp Ala Gly Asp Gln Glu Trp Glu Thr Cys Ala Arg  
 565 570 575

Glu Asn Asn Val Gly Trp Met Glu Met Ala Glu Pro Val Met Asn Leu  
 580 585 590

Tyr Thr Glu Thr Thr Asp Gly Ser Tyr Ile Glu Lys Lys Glu Thr Ala  
 595 600 605

Met Val Trp His Tyr Glu Asp Ala Asp Lys Asp Leu Gly Leu Glu Gln  
 610 615 620

Ala Lys Glu Leu Leu Asp His Leu Glu Asn Val Leu Ala Asn Glu Pro  
 625 630 635 640

Val Glu Val Lys Arg Gly Gln Tyr Ile Val Glu Val Lys Pro Gln Val  
 645 650 655

Pro His Gly Leu Pro Ser Cys Tyr Asp Ile His Arg His Arg Phe Val  
 660 665 670

Glu Ser Phe Asn Leu Asn Phe Phe Lys Tyr Glu Cys Asn Tyr Arg Gly  
 675 680 685

Ser Leu Lys Gly Ile Val Ala Glu Lys Ile Phe Ala Phe Met Ala Glu  
 690 695 700

Lys Gly Lys Gln Ala Asp Phe Val Leu Ser Val Gly Asp Asp Arg Ser  
 705 710 715 720

Asp Glu Asp Met Phe Val Ala Ile Gly Asp Gly Ile Lys Lys Gly Arg  
 725 730 735

Ile Thr Asn Asn Asn Ser Val Phe Thr Cys Val Val Gly Glu Lys Pro  
 740 745 750

Ser Ala Ala Glu Tyr Phe Leu Asp Glu Thr Lys Asp Val Ser Met Met  
755 760 765

Leu Glu Lys Leu Gly Cys Leu Ser Asn Gln Gly  
770 775

(2) INFORMATION FOR SEQ ID NO: 26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2130 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helianthus annuus*

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 171..2130
- (D) OTHER INFORMATION: /partial

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

GGATCCTGCG GTTTCATCAC ACAATATGAT ACTGTTACAT CTGATGCCCC TTCAGATGTC	60
CCAAATAGGT TGATTGTCGT ATCGAATCAG TTACCCATAA TCGCTAGGCT AAGACTAACG	120
ACAATGGAGG GTCCTTTTGG GATTTCACTT GGGACGAGAG TTCGATTAC ATG CAC	176
	Met His
	1
ATC AAA GAT GCA TTA CCC GCA GCC GTT GAG GTT TTC TAT GTT GGC GCA	224
Ile Lys Asp Ala Leu Pro Ala Ala Val Glu Val Phe Tyr Val Gly Ala	
5 10 15	
CTA AGG GCT GAC GTT GGC CCT ACC GAA CAA GAT GAC GTG TCA AAG ACA	272
Leu Arg Ala Asp Val Gly Pro Thr Glu Gln Asp Asp Val Ser Lys Thr	
20 25 30	
TTG CTC GAT AGG TTT AAT TGC GTT GCG GTT TTT GTC CCT ACT TCA AAA	320
Leu Leu Asp Arg Phe Asn Cys Val Ala Val Phe Val Pro Thr Ser Lys	
35 40 45 50	
TGG GAC CAA TAT TAT CAC TGC TTT TGT AAG CAG TAT TTG TGG CCG ATA	368
Trp Asp Gln Tyr Tyr His Cys Phe Cys Lys Gln Tyr Leu Trp Pro Ile	
55 60 65	
TTT CAT TAC AAG GTT CCC GCT TCT GAC GTC AAG AGT GTC CCG AAT AGT	416
Phe His Tyr Lys Val Pro Ala Ser Asp Val Lys Ser Val Pro Asn Ser	
70 75 80	

CGG GAT TCA TGG AAC GCT TAT GTT CAC GTG AAC AAA GAG TTT TCC CAG	464
Arg Asp Ser Trp Asn Ala Tyr Val His Val Asn Lys Glu Phe Ser Gln	
85 90 95	
AAG GTG ATG GAG GCA GTA ACC AAT GCT AGC AAT TAT GTA TGG ATA CAT	512
Lys Val Met Glu Ala Val Thr Asn Ala Ser Asn Tyr Val Trp Ile His	
100 105 110	
GAC TAC CAT TTA ATG ACG CTA CCG ACT TTC TTG AGG CGG GAT TTT TGT	560
Asp Tyr His Leu Met Thr Leu Pro Thr Phe Leu Arg Arg Asp Phe Cys	
115 120 125 130	
CGT TTT AAA ATC GGT TTT TTT CTG CAT AGC CCG TTT CCT TCC TCG GAG	608
Arg Phe Lys Ile Gly Phe Phe Leu His Ser Pro Phe Pro Ser Ser Glu	
135 140 145	
GTT TAC AAG ACC CTA CCA ATG AGA AAC GAG CTC TTG AAG GGT CTG TTA	656
Val Tyr Lys Thr Leu Pro Met Arg Asn Glu Leu Leu Lys Gly Leu Leu	
150 155 160	
AAT GCT GAT CTT ATC GGG TTC CAT ACA TAC GAT TAT GCC CGT CAT TTT	704
Asn Ala Asp Leu Ile Gly Phe His Thr Tyr Asp Tyr Ala Arg His Phe	
165 170 175	
CTA ACG TGT TGT AGT CGA ATG TTT GGT TTG GAT CAT CAG TTG AAA AGG	752
Leu Thr Cys Cys Ser Arg Met Phe Gly Leu Asp His Gln Leu Lys Arg	
180 185 190	
GGG TAC ATT TTC TTG GAA TAT AAT GGA AGG AGC ATT GAG ATC AAG ATA	800
Gly Tyr Ile Phe Leu Glu Tyr Asn Gly Arg Ser Ile Glu Ile Lys Ile	
195 200 205 210	
AAG GCG AGC GGG ATT CAT GTT GGT CGA ATG GAG TCG TAC TTG AGT CAG	848
Lys Ala Ser Gly Ile His Val Gly Arg Met Glu Ser Tyr Leu Ser Gln	
215 220 225	
CCC GAT ACA AGA TTA CAA GTT CAA GAA CTA AAA AAA CGT TTC GAA GGG	896
Pro Asp Thr Arg Leu Gln Val Gln Glu Leu Lys Lys Arg Phe Glu Gly	
230 235 240	
AAA ATC GTG CTA CTT GGA GTT GAT GAT TTG GAT ATA TTC AAA GGT GTG	944
Lys Ile Val Leu Leu Gly Val Asp Asp Leu Asp Ile Phe Lys Gly Val	
245 250 255	
AAC TTC AAG GTT TTA GCG TTG GAG AAG TTA CTT AAA TCA CAC CCG AGT	992
Asn Phe Lys Val Leu Ala Leu Glu Lys Leu Leu Lys Ser His Pro Ser	
260 265 270	
TGG CAA GGG CGT GTG GTT TTG GTG CAA ATC TTG AAT CCC GCT CGC GCG	1040
Trp Gln Gly Arg Val Val Leu Val Gln Ile Leu Asn Pro Ala Arg Ala	
275 280 285 290	
CGT TGC CAA GAC GTC GAT GAG ATC AAT GCC GAG ATA AGA ACA GTC TGT	1088
Arg Cys Gln Asp Val Asp Glu Ile Asn Ala Glu Ile Arg Thr Val Cys	
295 300 305	

GAA AGA ATC AAT AAC GAA CTG GGA AGC CCG GGA TAC CAG CCC GTT GTG	1136
Glu Arg Ile Asn Asn Glu Leu Gly Ser Pro Gly Tyr Gln Pro Val Val	
310 315 320	
TTA ATT GAT GGG CCC GTT TCG TTA AGT GAA AAA GCT GCT TAT TAT GCT	1184
Leu Ile Asp Gly Pro Val Ser Leu Ser Glu Lys Ala Ala Tyr Tyr Ala	
325 330 335	
ATC GCC GAT ATG GCA ATT GTT ACA CCG TTA CGT GAC GGC ATG AAT CTT	1232
Ile Ala Asp Met Ala Ile Val Thr Pro Leu Arg Asp Gly Met Asn Leu	
340 345 350	
ATC CCG TAC GAG TAC GTC GTT TCC CGA CAA AGT GTT AAT GAC CCA AAT	1280
Ile Pro Tyr Glu Tyr Val Val Ser Arg Gln Ser Val Asn Asp Pro Asn	
355 360 365 370	
CCC AAT ACT CCA AAA AAG AGC ATG CTA GTG GTC TCC GAG TTC ATC GGG	1328
Pro Asn Thr Pro Lys Lys Ser Met Leu Val Val Ser Glu Phe Ile Gly	
375 380 385	
TGT TCA CTA TCT TTA ACC GGG GCC ATA CGG GTC AAC CCA TGG GAT GAG	1376
Cys Ser Leu Ser Leu Thr Gly Ala Ile Arg Val Asn Pro Trp Asp Glu	
390 395 400	
TTG GAG ACA GCA GAA GCA TTA TAC GAC GCA CTC ATG GCT CCT GAT GAC	1424
Leu Glu Thr Ala Glu Ala Leu Tyr Asp Ala Leu Met Ala Pro Asp Asp	
405 410 415	
CAT AAA GAA ACC GCC CAC ATG AAA CAG TAT CAA TAC ATT ATC TCC CAT	1472
His Lys Glu Thr Ala His Met Lys Gln Tyr Gln Tyr Ile Ile Ser His	
420 425 430	
GAT GTA GCT AAC TGG GCT CGT AGC TTC TTT CAA GAT TTA GAG CAA GCG	1520
Asp Val Ala Asn Trp Ala Arg Ser Phe Phe Gln Asp Leu Glu Gln Ala	
435 440 445 450	
TGC ATC GAT CAT TCT CGT AAA CGA TGC ATG AAT TTA GGA TTT GGG TTA	1568
Cys Ile Asp His Ser Arg Lys Arg Cys Met Asn Leu Gly Phe Gly Leu	
455 460 465	
GAT ACT AGA GTC GTT CTT TTT GAT GAG AAG TTT AGC AAG TTG GAT ATA	1616
Asp Thr Arg Val Val Leu Phe Asp Glu Lys Phe Ser Lys Leu Asp Ile	
470 475 480	
GAT GTC TTG GAG AAT GCT TAT TCC ATG GCT CAA AAT CGG GCC ATA CTT	1664
Asp Val Leu Glu Asn Ala Tyr Ser Met Ala Gln Asn Arg Ala Ile Leu	
485 490 495	
TTG GAC TAT GAC GGC ACT GTT ACT CCA TCT ATC AGT AAA TCT CCA ACT	1712
Leu Asp Tyr Asp Gly Thr Val Thr Pro Ser Ile Ser Lys Ser Pro Thr	
500 505 510	
GAA GCT GTT ATC TCC ATG ATC AAC AAA CTG TGC AAT GAT CCA AAG AAC	1760
Glu Ala Val Ile Ser Met Ile Asn Lys Leu Cys Asn Asp Pro Lys Asn	
515 520 525 530	



ATG GTG TTC ATC GTT AGT GGA CGC AGT AGA GAA AAT CTT GGC AGT TGG	1808
Met Val Phe Ile Val Ser Gly Arg Ser Arg Glu Asn Leu Gly Ser Trp	
535 540 545	
TTC GGC GCG TGT GAG AAA CCC GCC ATT GCA GCT GAG CAC GGA TAC TTT	1856
Phe Gly Ala Cys Glu Lys Pro Ala Ile Ala Ala Glu His Gly Tyr Phe	
550 555 560	
ATA AGG TGG GCG GGT GAT CAA GAA TGG GAA ACG TGC GCA CGT GAG AAT	1904
Ile Arg Trp Ala Gly Asp Gln Glu Trp Glu Thr Cys Ala Arg Glu Asn	
565 570 575	
AAT GTC GGG TGG ATG GAA ATG GCT GAG CCG GTT ATG AAT CTT TAT ACA	1952
Asn Val Gly Trp Met Glu Met Ala Glu Pro Val Met Asn Leu Tyr Thr	
580 585 590	
GAA ACT ACT GAC GGT TCG TAT ATT GAA AAG AAA GAA ACT GCA ATG GTT	2000
Glu Thr Thr Asp Gly Ser Tyr Ile Glu Lys Lys Glu Thr Ala Met Val	
595 600 605 610	
TGG CAC TAT GAA GAT GCT GAT AAA GAT CTT GGG TTG GAG CAG GCT AAG	2048
Trp His Tyr Glu Asp Ala Asp Lys Asp Leu Gly Leu Glu Gln Ala Lys	
615 620 625	
GAA CTG TTG GAC CAT CTT GAA AAC GTG CTC GCT AAT GAG CCC GTT GAA	2096
Glu Leu Leu Asp His Leu Glu Asn Val Leu Ala Asn Glu Pro Val Glu	
630 635 640	
GTG AAA CGA GGT CAA TAC ATT GTA GAA GTT AAA C	2130
Val Lys Arg Gly Gln Tyr Ile Val Glu Val Lys	
645 650	

(2) INFORMATION FOR SEQ ID NO: 27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 653 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

Met His Ile Lys Asp Ala Leu Pro Ala Ala Val Glu Val Phe Tyr Val	
1 5 10 15	
Gly Ala Leu Arg Ala Asp Val Gly Pro Thr Glu Gln Asp Asp Val Ser	
20 25 30	
Lys Thr Leu Leu Asp Arg Phe Asn Cys Val Ala Val Phe Val Pro Thr	
35 40 45	
Ser Lys Trp Asp Gln Tyr Tyr His Cys Phe Cys Lys Gln Tyr Leu Trp	
50 55 60	
Pro Ile Phe His Tyr Lys Val Pro Ala Ser Asp Val Lys Ser Val Pro	
65 70 75 80	

Asn	Ser	Arg	Asp	Ser	Trp	Asn	Ala	Tyr	Val	His	Val	Asn	Lys	Glu	Phe	85	90	95	
Ser	Gln	Lys	Val	Met	Glu	Ala	Val	Thr	Asn	Ala	Ser	Asn	Tyr	Val	Trp	100	105	110	
Ile	His	Asp	Tyr	His	Leu	Met	Thr	Leu	Pro	Thr	Phe	Leu	Arg	Arg	Asp	115	120	125	
Phe	Cys	Arg	Phe	Lys	Ile	Gly	Phe	Phe	Leu	His	Ser	Pro	Phe	Pro	Ser	130	135	140	
Ser	Glu	Val	Tyr	Lys	Thr	Leu	Pro	Met	Arg	Asn	Glu	Leu	Leu	Lys	Gly	145	150	155	160
Leu	Leu	Asn	Ala	Asp	Leu	Ile	Gly	Phe	His	Thr	Tyr	Asp	Tyr	Ala	Arg	165	170	175	
His	Phe	Leu	Thr	Cys	Cys	Ser	Arg	Met	Phe	Gly	Leu	Asp	His	Gln	Leu	180	185	190	
Lys	Arg	Gly	Tyr	Ile	Phe	Leu	Glu	Tyr	Asn	Gly	Arg	Ser	Ile	Glu	Ile	195	200	205	
Lys	Ile	Lys	Ala	Ser	Gly	Ile	His	Val	Gly	Arg	Met	Glu	Ser	Tyr	Leu	210	215	220	
Ser	Gln	Pro	Asp	Thr	Arg	Leu	Gln	Val	Gln	Glu	Leu	Lys	Lys	Arg	Phe	225	230	235	240
Glu	Gly	Lys	Ile	Val	Leu	Leu	Gly	Val	Asp	Asp	Leu	Asp	Ile	Phe	Lys	245	250	255	
Gly	Val	Asn	Phe	Lys	Val	Leu	Ala	Leu	Glu	Lys	Leu	Leu	Lys	Ser	His	260	265	270	
Pro	Ser	Trp	Gln	Gly	Arg	Val	Val	Leu	Val	Gln	Ile	Leu	Asn	Pro	Ala	275	280	285	
Arg	Ala	Arg	Cys	Gln	Asp	Val	Asp	Glu	Ile	Asn	Ala	Glu	Ile	Arg	Thr	290	295	300	
Val	Cys	Glu	Arg	Ile	Asn	Asn	Glu	Leu	Gly	Ser	Pro	Gly	Tyr	Gln	Pro	305	310	315	320
Val	Val	Leu	Ile	Asp	Gly	Pro	Val	Ser	Leu	Ser	Glu	Lys	Ala	Ala	Tyr	325	330	335	
Tyr	Ala	Ile	Ala	Asp	Met	Ala	Ile	Val	Thr	Pro	Leu	Arg	Asp	Gly	Met	340	345	350	
Asn	Leu	Ile	Pro	Tyr	Glu	Tyr	Val	Val	Ser	Arg	Gln	Ser	Val	Asn	Asp	355	360	365	
Pro	Asn	Pro	Asn	Thr	Pro	Lys	Lys	Ser	Met	Leu	Val	Val	Ser	Glu	Phe	370	375	380	

Ile Gly Cys Ser Leu Ser Leu Thr Gly Ala Ile Arg Val Asn Pro Trp  
 385 390 395 400  
 Asp Glu Leu Glu Thr Ala Glu Ala Leu Tyr Asp Ala Leu Met Ala Pro  
 405 410 415  
 Asp Asp His Lys Glu Thr Ala His Met Lys Gln Tyr Gln Tyr Ile Ile  
 420 425 430  
 Ser His Asp Val Ala Asn Trp Ala Arg Ser Phe Phe Gln Asp Leu Glu  
 435 440 445  
 Gln Ala Cys Ile Asp His Ser Arg Lys Arg Cys Met Asn Leu Gly Phe  
 450 455 460  
 Gly Leu Asp Thr Arg Val Val Leu Phe Asp Glu Lys Phe Ser Lys Leu  
 465 470 475 480  
 Asp Ile Asp Val Leu Glu Asn Ala Tyr Ser Met Ala Gln Asn Arg Ala  
 485 490 495  
 Ile Leu Leu Asp Tyr Asp Gly Thr Val Thr Pro Ser Ile Ser Lys Ser  
 500 505 510  
 Pro Thr Glu Ala Val Ile Ser Met Ile Asn Lys Leu Cys Asn Asp Pro  
 515 520 525  
 Lys Asn Met Val Phe Ile Val Ser Gly Arg Ser Arg Glu Asn Leu Gly  
 530 535 540  
 Ser Trp Phe Gly Ala Cys Glu Lys Pro Ala Ile Ala Ala Glu His Gly  
 545 550 555 560  
 Tyr Phe Ile Arg Trp Ala Gly Asp Gln Glu Trp Glu Thr Cys Ala Arg  
 565 570 575  
 Glu Asn Asn Val Gly Trp Met Glu Met Ala Glu Pro Val Met Asn Leu  
 580 585 590  
 Tyr Thr Glu Thr Thr Asp Gly Ser Tyr Ile Glu Lys Lys Glu Thr Ala  
 595 600 605  
 Met Val Trp His Tyr Glu Asp Ala Asp Lys Asp Leu Gly Leu Glu Gln  
 610 615 620  
 Ala Lys Glu Leu Leu Asp His Leu Glu Asn Val Leu Ala Asn Glu Pro  
 625 630 635 640  
 Val Glu Val Lys Arg Gly Gln Tyr Ile Val Glu Val Lys  
 645 650

(2) INFORMATION FOR SEQ ID NO: 28:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 390 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Helianthus annuus

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 3..258

(D) OTHER INFORMATION: /partial

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

TT GCA GAG AAG ATT TTT GCG TTC ATG GCT GAA AAG GGA AAA CAG GCT	47
Ala Glu Lys Ile Phe Ala Phe Met Ala Glu Lys Gly Lys Gln Ala	
1 5 10 15	
GAT TTC GTG TTG AGC GTT GGA GAT GAT AGA AGT GAT GAA GAC ATG TTT	95
Asp Phe Val Leu Ser Val Gly Asp Asp Arg Ser Asp Glu Asp Met Phe	
20 25 30	
GTG GCC ATT GGG GAT GGA ATA AAA AAG GGT CGG ATA ACT AAC AAC AAT	143
Val Ala Ile Gly Asp Gly Ile Lys Lys Gly Arg Ile Thr Asn Asn Asn	
35 40 45	
TCA GTG TTT ACA TGC GTA GTG GGA GAG AAA CCG AGT GCA GCT GAG TAC	191
Ser Val Phe Thr Cys Val Val Gly Glu Lys Pro Ser Ala Ala Glu Tyr	
50 55 60	
TTT TTA GAC GAG ACG AAA GAT GTT TCA ATG ATG CTC GAG AAG CTC GGG	239
Phe Leu Asp Glu Thr Lys Asp Val Ser Met Met Leu Glu Lys Leu Gly	
65 70 75	
TGT CTC AGC AAC CAA GGA T GATGATCCGG AAGCTTCTCG TGATCTTTAT	288
Cys Leu Ser Asn Gln Gly	
80 85	
GAGTTAAAAG TTTTCGACTT TTTCTTCATC AAGATTCATG GGAAAGTTGT TCAATATGAA	348
CTTGTGTTTC TTGGTTCTGG ATTTTAGGGA GTCTATGGAT CC	390

(2) INFORMATION FOR SEQ ID NO: 29:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 85 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

Ala Glu Lys Ile Phe Ala Phe Met Ala Glu Lys Gly Lys Gln Ala Asp  
 1 5 10 15  
 Phe Val Leu Ser Val Gly Asp Asp Arg Ser Asp Glu Asp Met Phe Val  
 20 25 30  
 Ala Ile Gly Asp Gly Ile Lys Lys Gly Arg Ile Thr Asn Asn Asn Ser  
 35 40 45  
 Val Phe Thr Cys Val Val Gly Glu Lys Pro Ser Ala Ala Glu Tyr Phe  
 50 55 60  
 Leu Asp Glu Thr Lys Asp Val Ser Met Met Leu Glu Lys Leu Gly Cys  
 65 70 75 80  
 Leu Ser Asn Gln Gly  
 85

(2) INFORMATION FOR SEQ ID NO: 30:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (ix) FEATURES:
  - (A) NAME/KEY: modified base, Inosine
  - (B) LOCATION: 4
  - (C) OTHER INFORMATION: /note= N is Inosine
- (ix) FEATURES:
  - (A) NAME/KEY: modified base, Inosine
  - (B) LOCATION: 10
  - (C) OTHER INFORMATION: /note= N is Inosine
- (ix) FEATURES:
  - (A) NAME/KEY: modified base, Inosine
  - (B) LOCATION: 13
  - (C) OTHER INFORMATION: /note= N is Inosine
- (ix) FEATURES:
  - (A) NAME/KEY: modified base, Inosine
  - (B) LOCATION: 19
  - (C) OTHER INFORMATION: /note= N is Inosine
- (ix) FEATURES:
  - (A) NAME/KEY: modified base, Inosine
  - (B) LOCATION: 22
  - (C) OTHER INFORMATION: /note= N is Inosine

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:  
CCANGGRTTN ACNCKDNTNG CNCC

24

(2) INFORMATION FOR SEQ ID NO: 31:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 23 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (ix) FEATURES:
  - (A) NAME/KEY: modified base, Inosine
  - (B) LOCATION: 6
  - (C) OTHER INFORMATION: /note= A is Inosine
- (ix) FEATURES:
  - (A) NAME/KEY: modified base, Inosine
  - (B) LOCATION: 9
  - (C) OTHER INFORMATION: /note= A is Inosine
- (ix) FEATURES:
  - (A) NAME/KEY: modified base, Inosine
  - (B) LOCATION: 12
  - (C) OTHER INFORMATION: /note= A is Inosine
- (ix) FEATURES:
  - (A) NAME/KEY: modified base, Inosine
  - (B) LOCATION: 18
  - (C) OTHER INFORMATION: /note= A is Inosine
- (ix) FEATURES:
  - (A) NAME/KEY: modified base, Inosine
  - (B) LOCATION: 21
  - (C) OTHER INFORMATION: /note= A is Inosine

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:  
ATHGTNGTNW SNAAYMRNYT NCC

23

(2) INFORMATION FOR SEQ ID NO: 32:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURES:

(A) NAME/KEY: modified base, Inosine

(B) LOCATION: 3

(C) OTHER INFORMATION: /note= N is Inosine

(ix) FEATURES:

(A) NAME/KEY: modified base, Inosine

(B) LOCATION: 9

(C) OTHER INFORMATION: /note= N is Inosine

(ix) FEATURES:

(A) NAME/KEY: modified base, Inosine

(B) LOCATION: 12

(C) OTHER INFORMATION: /note= N is Inosine

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:

YTNTGGCCNA TNTTYCAYTA

20

(2) INFORMATION FOR SEQ ID NO: 33:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURES:

(A) NAME/KEY: modified base, Inosine

(B) LOCATION: 6

(C) OTHER INFORMATION: /note= N is Inosine

(ix) FEATURES:

(A) NAME/KEY: modified base, Inosine

(B) LOCATION: 9

(C) OTHER INFORMATION: /note= N is Inosine

(ix) FEATURES:

(A) NAME/KEY: modified base, Inosine

(B) LOCATION: 18

(C) OTHER INFORMATION: /note= N is Inosine

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 33:

TGRTCNARNA RYTCYTTNGC

20

(2) INFORMATION FOR SEQ ID NO: 34:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (ix) FEATURES:
  - (A) NAME/KEY: modified base, Inosine
  - (B) LOCATION: 6
  - (C) OTHER INFORMATION: /note= N is Inosine
- (ix) FEATURES:
  - (A) NAME/KEY: modified base, Inosine
  - (B) LOCATION: 18
  - (C) OTHER INFORMATION: /note= N is Inosine
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 34:

TCRTCNGTRA ARTCRTCNCC

20

(2) INFORMATION FOR SEQ ID NO: 35:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (ix) FEATURES:
  - (A) NAME/KEY: modified base, Inosine
  - (B) LOCATION: 15
  - (C) OTHER INFORMATION: /note= N is Inosine
- (ix) FEATURES:
  - (A) NAME/KEY: modified base, Inosine
  - (B) LOCATION: 18
  - (C) OTHER INFORMATION: /note= N is Inosine
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 35:

TTYGAYTAYG AYGGNACNYT

20



(2) INFORMATION FOR SEQ ID NO: 36:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (ix) FEATURES:
  - (A) NAME/KEY: modified base, Inosine
  - (B) LOCATION: 3
  - (C) OTHER INFORMATION: /note= N is Inosine
- (ix) FEATURES:
  - (A) NAME/KEY: modified base, Inosine
  - (B) LOCATION: 6
  - (C) OTHER INFORMATION: /note= N is Inosine
- (ix) FEATURES:
  - (A) NAME/KEY: modified base, Inosine
  - (B) LOCATION: 12
  - (C) OTHER INFORMATION: /note= N is Inosine
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 36:

GGNYTNWBNG CNGARCA YGG

20

(2) INFORMATION FOR SEQ ID NO: 37:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (ix) FEATURES:
  - (A) NAME/KEY: modified base, Inosine
  - (B) LOCATION: 3
  - (C) OTHER INFORMATION: /note= N is Inosine
- (ix) FEATURES:
  - (A) NAME/KEY: modified base, Inosine
  - (B) LOCATION: 6
  - (C) OTHER INFORMATION: /note= N is Inosine

(ix) FEATURES:  
    (A) NAME/KEY: modified base, Inosine  
    (B) LOCATION: 12  
    (C) OTHER INFORMATION: /note= N is Inosine

(ix) FEATURES:  
    (A) NAME/KEY: modified base, Inosine  
    (B) LOCATION: 15  
    (C) OTHER INFORMATION: /note= N is Inosine

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 37:

ATNGCNAARC CNGTNATGAA

20

(2) INFORMATION FOR SEQ ID NO: 38:

(i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 20 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURES:  
    (A) NAME/KEY: modified base, Inosine  
    (B) LOCATION: 3  
    (C) OTHER INFORMATION: /note= N is Inosine

(ix) FEATURES:  
    (A) NAME/KEY: modified base, Inosine  
    (B) LOCATION: 6  
    (C) OTHER INFORMATION: /note= N is Inosine

(ix) FEATURES:  
    (A) NAME/KEY: modified base, Inosine  
    (B) LOCATION: 12  
    (C) OTHER INFORMATION: /note= N is Inosine

(ix) FEATURES:  
    (A) NAME/KEY: modified base, Inosine  
    (B) LOCATION: 18  
    (C) OTHER INFORMATION: /note= N is Inosine

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 38:

CCNACNGTRC ANGCRANAC

20

(2) INFORMATION FOR SEQ ID NO: 39:

(i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 2982 base pairs

(B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Arabidopsis thaliana

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 64..2889

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 39:

ATAAACTTCC TCGCGGCCGC CAGTGTGAGT AATTTAGTTT TGTTCTGTT TTGGTGTGAG	60
CGT ATG CCT GGA AAT AAG TAC AAC TGC AGT TCT TCT CAT ATC CCA CTC	108
Met Pro Gly Asn Lys Tyr Asn Cys Ser Ser Ser His Ile Pro Leu	
1 5 10 15	
TCT CGA ACA GAA CGC CTC TTG AGA GAT AGA GAG CTT AGA GAG AAG AGG	156
Ser Arg Thr Glu Arg Leu Leu Arg Asp Arg Glu Leu Arg Glu Lys Arg	
20 25 30	
AAG AGC AAC CGA GCT CGT AAT CCT AAT GAC GTT GCT GGC AGT TCC GAG	204
Lys Ser Asn Arg Ala Arg Asn Pro Asn Asp Val Ala Gly Ser Ser Glu	
35 40 45	
AAC TCT GAG AAT GAC TTG CGT TTA GAA GGT GAC AGT TCA AGG CAG TAT	252
Asn Ser Glu Asn Asp Leu Arg Leu Glu Gly Asp Ser Ser Arg Gln Tyr	
50 55 60	
GTT GAA CAG TAC TTG GAA GGG GCT GCT GCT GCA ATG GCG CAC GAT GAT	300
Val Glu Gln Tyr Leu Glu Gly Ala Ala Ala Ala Met Ala His Asp Asp	
65 70 75	
GCG TGT GAG AGG CAA GAA GTT AGG CCT TAT AAT AGG CAA CGA CTA CTT	348
Ala Cys Glu Arg Gln Glu Val Arg Pro Tyr Asn Arg Gln Arg Leu Leu	
80 85 90 95	
GTA GTG GCT AAC AGG CTC CCA GTT TCT CCC GTG AGA AGA GGT GAA GAT	396
Val Val Ala Asn Arg Leu Pro Val Ser Pro Val Arg Arg Gly Glu Asp	
100 105 110	
TCA TGG TCT CTT GAG ATC AGT GCT GGT GGT CTA GTC AGT GCT CTC TTA	444
Ser Trp Ser Leu Glu Ile Ser Ala Gly Gly Leu Val Ser Ala Leu Leu	
115 120 125	
GGT GTA AAG GAA TTT GAG GCC AGA TGG ATA GGA TGG GCT GGA GTT AAT	492
Gly Val Lys Glu Phe Glu Ala Arg Trp Ile Gly Trp Ala Gly Val Asn	
130 135 140	

GTG CCT GAT GAG GTT GGA CAG AAG GCA CTT AGC AAA GCT TTG GCT GAG Val Pro Asp Glu Val Gly Gln Lys Ala Leu Ser Lys Ala Leu Ala Glu 145 150 155	540
AAG AGG TGT ATT CCC GTG TTC CTT GAT GAA GAG ATT GTT CAT CAG TAC Lys Arg Cys Ile Pro Val Phe Leu Asp Glu Glu Ile Val His Gln Tyr 160 165 170 175	588
TAT AAT GGT TAC TGC AAC AAT ATT CTG TGG CCT CTG TTT CAC TAC CTT Tyr Asn Gly Tyr Cys Asn Asn Ile Leu Trp Pro Leu Phe His Tyr Leu 180 185 190	636
GGA CTT CCG CAA GAA GAT CGG CTT GCC ACA ACC AGA AGC TTT CAG TCC Gly Leu Pro Gln Glu Asp Arg Leu Ala Thr Thr Arg Ser Phe Gln Ser 195 200 205	684
CAA TTT GCT GCA TAC AAG AAG GCA AAC CAA ATG TTC GCT GAT GTT GTA Gln Phe Ala Ala Tyr Lys Lys Ala Asn Gln Met Phe Ala Asp Val Val 210 215 220	732
AAT GAG CAC TAT GAA GAG GGA GAT GTC GTC TGG TGC CAT GAC TAT CAT Asn Glu His Tyr Glu Glu Gly Asp Val Val Trp Cys His Asp Tyr His 225 230 235	780
CTT ATG TTC CTT CCT AAA TGC CTT AAG GAG TAC AAC AGT AAG ATG AAA Leu Met Phe Leu Pro Lys Cys Leu Lys Glu Tyr Asn Ser Lys Met Lys 240 245 250 255	828
GTT GGA TGG TTT CTC CAT ACA CCA TTC CCT TCG TCT GAG ATA CAC AGG Val Gly Trp Phe Leu His Thr Pro Phe Pro Ser Ser Glu Ile His Arg 260 265 270	876
ACA CTT CCA TCA CGA TCA GAG CTC CTT CGG TCA GTT CTT GCT GCT GAT Thr Leu Pro Ser Arg Ser Glu Leu Leu Arg Ser Val Leu Ala Ala Asp 275 280 285	924
TTA GTT GGC TTC CAT ACA TAT GAC TAT GCA AGG CAC TTT GTG AGT GCG Leu Val Gly Phe His Thr Tyr Asp Tyr Ala Arg His Phe Val Ser Ala 290 295 300	972
TGC ACT CGT ATT CTT GGA CTT GAA GGA ACA CCT GAG GGA GTT GAG GAT Cys Thr Arg Ile Leu Gly Leu Glu Gly Thr Pro Glu Gly Val Glu Asp 305 310 315	1020
CAA GGC AGG CTC ACT CGT GTA GCT GCT TTT CCA ATT GGC ATA GAT TCT Gln Gly Arg Leu Thr Arg Val Ala Ala Phe Pro Ile Gly Ile Asp Ser 320 325 330 335	1068
GAT CGG TTT ATA CGA GCA CTT GAG GTC CCC GAA GTC AAA CAA CAC ATG Asp Arg Phe Ile Arg Ala Leu Glu Val Pro Glu Val Lys Gln His Met 340 345 350	1116
AAG GAA TTG AAA GAA AGA TTT ACT GAC AGA AAG GTG ATG TTA GGT GTT Lys Glu Leu Lys Glu Arg Phe Thr Asp Arg Lys Val Met Leu Gly Val 355 360 365	1164

GAT CGT CTT GAC ATG ATC AAA GGG ATT CCA CAA AAG ATT CTG GCA TTC Asp Arg Leu Asp Met Ile Lys Gly Ile Pro Gln Lys Ile Leu Ala Phe 370 375 380	1212
GAA AAA TTT CTC GAG GAA AAT GCA AAC TGG CGT GAT AAA GTG GTC TTA Glu Lys Phe Leu Glu Glu Asn Ala Asn Trp Arg Asp Lys Val Val Leu 385 390 395	1260
TTG AAA ATT GCG GTG CCA ACA AGA CCT GAC GTT CCT GAG TAT CAA ACA Leu Lys Ile Ala Val Pro Thr Arg Pro Asp Val Pro Glu Tyr Gln Thr 400 405 410 415	1308
CTC ACA AGC CAA GTT CAT GAA ATT GTT GGC CGC ATT ATT GGT CGT CTC Leu Thr Ser Gln Val His Glu Ile Val Gly Arg Ile Ile Gly Arg Leu 420 425 430	1356
GGG ACA CTG ACT GCA GTT CCA ATA CAT CAT CTG GAT CGG TCT CTG GAC Gly Thr Leu Thr Ala Val Pro Ile His His Leu Asp Arg Ser Leu Asp 435 440 445	1404
TTT CAT GCT TTA TGT GCA CTT TAT GCC GTC ACA GAT GTT GCG CTT GTA Phe His Ala Leu Cys Ala Leu Tyr Ala Val Thr Asp Val Ala Leu Val 450 455 460	1452
ACA TCT TTG AGA GAT GGG ATG AAT CTT GTC AGT TAT GAG TTT GTT GCT Thr Ser Leu Arg Asp Gly Met Asn Leu Val Ser Tyr Glu Phe Val Ala 465 470 475	1500
TGC CAA GAG GCC AAA AAG GGC GTC CTC ATT CTC AGT GAA TTT GCA GGT Cys Gln Glu Ala Lys Lys Gly Val Leu Ile Leu Ser Glu Phe Ala Gly 480 485 490 495	1548
GCT GCA CAG TCT CTG GGT GCT GGA GCT ATT CTT GTG AAT CCT TGG AAC Ala Ala Gln Ser Leu Gly Ala Gly Ala Ile Leu Val Asn Pro Trp Asn 500 505 510	1596
ATC ACA GAA GTT GCT GCC TCC ATT GGA CAA GCC CTA AAC ATG ACA GCT Ile Thr Glu Val Ala Ala Ser Ile Gly Gln Ala Leu Asn Met Thr Ala 515 520 525	1644
GAA GAA AGA GAG AAA AGA CAT CGC CAT AAT TTT CAT CAT GTC AAA ACT Glu Glu Arg Glu Lys Arg His Arg His Asn Phe His His Val Lys Thr 530 535 540	1692
CAC ACT GCT CAA GAA TGG GCT GAA ACT TTT GTC AGT GAA CTA AAT GAC His Thr Ala Gln Glu Trp Ala Glu Thr Phe Val Ser Glu Leu Asn Asp 545 550 555	1740
ACT GTA ATT GAG GCG CAA CTA CGA ATT AGT AAA GTC CCA CCA GAG CTT Thr Val Ile Glu Ala Gln Leu Arg Ile Ser Lys Val Pro Pro Glu Leu 560 565 570 575	1788
CCA CAG CAT GAT GCA ATT CAA CGG TAT TCA AAG TCC AAC AAC AGG CTT Pro Gln His Asp Ala Ile Gln Arg Tyr Ser Lys Ser Asn Asn Arg Leu 580 585 590	1836

CTA ATC CTG GGT TTC AAT GCA ACA TTG ACT GAA CCA GTG GAT AAT CAA	1884
Leu Ile Leu Gly Phe Asn Ala Thr Leu Thr Glu Pro Val Asp Asn Gln	
595 600 605	
GGG AGA AGA GGT GAT CAA ATA AAG GAG ATG GAT CTT AAT CTA CAC CCT	1932
Gly Arg Arg Gly Asp Gln Ile Lys Glu Met Asp Leu Asn Leu His Pro	
610 615 620	
GAG CTT AAA GGG CCC TTA AAG GCA TTA TGC AGT GAT CCA AGT ACA ACC	1980
Glu Leu Lys Gly Pro Leu Lys Ala Leu Cys Ser Asp Pro Ser Thr Thr	
625 630 635	
ATA GTT GTT CTG AGC GGA AGC AGC AGA AGT GTT TTG GAC AAA AAC TTT	2028
Ile Val Val Leu Ser Gly Ser Ser Arg Ser Val Leu Asp Lys Asn Phe	
640 645 650 655	
GGA GAG TAT GAC ATG TGG CTG GCA GCA GAA AAT GGG ATG TTC CTA AGG	2076
Gly Glu Tyr Asp Met Trp Leu Ala Ala Glu Asn Gly Met Phe Leu Arg	
660 665 670	
CTT ACG AAT GGA GAG TGG ATG ACT ACA ATG CCA GAA CAC TTG AAC ATG	2124
Leu Thr Asn Gly Glu Trp Met Thr Thr Met Pro Glu His Leu Asn Met	
675 680 685	
GAA TGG GTT GAT AGC GTA AAG CAT GTT TTC AAG TAC TTC ACT GAG AGA	2172
Glu Trp Val Asp Ser Val Lys His Val Phe Lys Tyr Phe Thr Glu Arg	
690 695 700	
ACT CCC AGG TCA CAC TTT GAA ACT CGC GAT ACT TCG CTT ATT TGG AAC	2220
Thr Pro Arg Ser His Phe Glu Thr Arg Asp Thr Ser Leu Ile Trp Asn	
705 710 715	
TAC AAA TAT GCA GAT ATC GAA TTC GGG AGA CTT CAA GCA AGA GAT TTG	2268
Tyr Lys Tyr Ala Asp Ile Glu Phe Gly Arg Leu Gln Ala Arg Asp Leu	
720 725 730 735	
TTA CAA CAC TTA TGG ACA GGT CCA ATC TCT AAT GCA TCA GTT GAT GTT	2316
Leu Gln His Leu Trp Thr Gly Pro Ile Ser Asn Ala Ser Val Asp Val	
740 745 750	
GTC CAA GGA AGC CGC TCT GTG GAA GTC CGT GCA GTT GGT GTC ACA AAG	2364
Val Gln Gly Ser Arg Ser Val Glu Val Arg Ala Val Gly Val Thr Lys	
755 760 765	
GGA GCT GCA ATT GAT CGT ATT CTA GGA GAG ATA GTG CAT AGC AAG TCG	2412
Gly Ala Ala Ile Asp Arg Ile Leu Gly Glu Ile Val His Ser Lys Ser	
770 775 780	
ATG ACT ACA CCA ATC GAT TAC GTC TTG TGC ATT GGT CAT TTC TTG GGG	2460
Met Thr Thr Pro Ile Asp Tyr Val Leu Cys Ile Gly His Phe Leu Gly	
785 790 795	
AAG GAC GAA GAT GTT TAC ACT TTC TTC GAA CCA GAA CTT CCA TCC GAC	2508
Lys Asp Glu Asp Val Tyr Thr Phe Phe Glu Pro Glu Leu Pro Ser Asp	
800 805 810 815	

ATG CCA GCC ATT GCA CGA TCC AGA CCA TCA TCT GAC AGT GGA GCC AAG	2556
Met Pro Ala Ile Ala Arg' Ser Arg Pro Ser Ser Asp Ser Gly Ala Lys	
820 825 830	
TCA TCA TCA GGA GAC CGA AGA CCA CCT TCA AAG TCG ACA CAT AAC AAC	2604
Ser Ser Ser Gly Asp Arg Arg Pro Pro Ser Lys Ser Thr His Asn Asn	
835 840 845	
AAC AAA AGT GGA TCA AAA TCC TCA TCA TCC TCT AAC TCT AAC AAC AAC	2652
Asn Lys Ser Gly Ser Lys Ser Ser Ser Ser Asn Ser Asn Asn Asn	
850 855 860	
AAC AAG TCC TCA CAG AGA TCT CTT CAG TCA GAG AGA AAA AGT GGA TCC	2700
Asn Lys Ser Ser Gln Arg Ser Leu Gln Ser Glu Arg Lys Ser Gly Ser	
865 870 875	
AAC CAT AGC TTA GGA AAC TCA AGA CGT CCT TCA CCA GAG AAG ATC TCA	2748
Asn His Ser Leu Gly Asn Ser Arg Arg Pro Ser Pro Glu Lys Ile Ser	
880 885 890 895	
TGG AAT GTG CTT GAC CTC AAA GGA GAG AAC TAC TTC TCT TGC GCT GTG	2796
Trp Asn Val Leu Asp Leu Lys Gly Glu Asn Tyr Phe Ser Cys Ala Val	
900 905 910	
GGT CGT ACT CGC ACC AAT GCT AGA TAT CTC CTT GGC TCA CCT GAC GAC	2844
Gly Arg Thr Arg Thr Asn Ala Arg Tyr Leu Leu Gly Ser Pro Asp Asp	
915 920 925	
GTC GTT TGC TTC CTT GAG AAG CTC GCT GAC ACC ACT TCC TCA CCT TAA	2892
Val Val Cys Phe Leu Glu Lys Leu Ala Asp Thr Thr Ser Ser Pro	
930 935 940	
TATCCCGAGA CAGTGTCAAG TGAGTTCATG TAACCCAATA AAAACTATTG TTTTGTAACA	2952
AAAAGCAGCC ATTACCAGAC TCTTTAGTGG	2982

(2) INFORMATION FOR SEQ ID NO: 40:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 942 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 40:

Met Pro Gly Asn Lys Tyr Asn Cys Ser Ser Ser His Ile Pro Leu Ser	
1 5 10 15	
Arg Thr Glu Arg Leu Leu Arg Asp Arg Glu Leu Arg Glu Lys Arg Lys	
20 25 30	
Ser Asn Arg Ala Arg Asn Pro Asn Asp Val Ala Gly Ser Ser Glu Asn	
35 40 45	

Ser	Glu	Asn	Asp	Leu	Arg	Leu	Glu	Gly	Asp	Ser	Ser	Arg	Gln	Tyr	Val	50	55	60	
Glu	Gln	Tyr	Leu	Glu	Gly	Ala	Ala	Ala	Ala	Met	Ala	His	Asp	Asp	Ala	65	70	75	80
Cys	Glu	Arg	Gln	Glu	Val	Arg	Pro	Tyr	Asn	Arg	Gln	Arg	Leu	Leu	Val	85	90	95	
Val	Ala	Asn	Arg	Leu	Pro	Val	Ser	Pro	Val	Arg	Arg	Gly	Glu	Asp	Ser	100	105	110	
Trp	Ser	Leu	Glu	Ile	Ser	Ala	Gly	Gly	Leu	Val	Ser	Ala	Leu	Leu	Gly	115	120	125	
Val	Lys	Glu	Phe	Glu	Ala	Arg	Trp	Ile	Gly	Trp	Ala	Gly	Val	Asn	Val	130	135	140	
Pro	Asp	Glu	Val	Gly	Gln	Lys	Ala	Leu	Ser	Lys	Ala	Leu	Ala	Glu	Lys	145	150	155	160
Arg	Cys	Ile	Pro	Val	Phe	Leu	Asp	Glu	Glu	Ile	Val	His	Gln	Tyr	Tyr	165	170	175	
Asn	Gly	Tyr	Cys	Asn	Asn	Ile	Leu	Trp	Pro	Leu	Phe	His	Tyr	Leu	Gly	180	185	190	
Leu	Pro	Gln	Glu	Asp	Arg	Leu	Ala	Thr	Thr	Arg	Ser	Phe	Gln	Ser	Gln	195	200	205	
Phe	Ala	Ala	Tyr	Lys	Lys	Ala	Asn	Gln	Met	Phe	Ala	Asp	Val	Val	Asn	210	215	220	
Glu	His	Tyr	Glu	Glu	Gly	Asp	Val	Val	Trp	Cys	His	Asp	Tyr	His	Leu	225	230	235	240
Met	Phe	Leu	Pro	Lys	Cys	Leu	Lys	Glu	Tyr	Asn	Ser	Lys	Met	Lys	Val	245	250	255	
Gly	Trp	Phe	Leu	His	Thr	Pro	Phe	Pro	Ser	Ser	Glu	Ile	His	Arg	Thr	260	265	270	
Leu	Pro	Ser	Arg	Ser	Glu	Leu	Leu	Arg	Ser	Val	Leu	Ala	Ala	Asp	Leu	275	280	285	
Val	Gly	Phe	His	Thr	Tyr	Asp	Tyr	Ala	Arg	His	Phe	Val	Ser	Ala	Cys	290	295	300	
Thr	Arg	Ile	Leu	Gly	Leu	Glu	Gly	Thr	Pro	Glu	Gly	Val	Glu	Asp	Gln	305	310	315	320
Gly	Arg	Leu	Thr	Arg	Val	Ala	Ala	Phe	Pro	Ile	Gly	Ile	Asp	Ser	Asp	325	330	335	
Arg	Phe	Ile	Arg	Ala	Leu	Glu	Val	Pro	Glu	Val	Lys	Gln	His	Met	Lys	340	345	350	



Glu Leu Lys Glu Arg Phe Thr Asp Arg Lys Val Met Leu Gly Val Asp  
 355 360 365  
 Arg Leu Asp Met Ile Lys Gly Ile Pro Gln Lys Ile Leu Ala Phe Glu  
 370 375 380  
 Lys Phe Leu Glu Glu Asn Ala Asn Trp Arg Asp Lys Val Val Leu Leu  
 385 390 395 400  
 Lys Ile Ala Val Pro Thr Arg Pro Asp Val Pro Glu Tyr Gln Thr Leu  
 405 410 415  
 Thr Ser Gln Val His Glu Ile Val Gly Arg Ile Ile Gly Arg Leu Gly  
 420 425 430  
 Thr Leu Thr Ala Val Pro Ile His His Leu Asp Arg Ser Leu Asp Phe  
 435 440 445  
 His Ala Leu Cys Ala Leu Tyr Ala Val Thr Asp Val Ala Leu Val Thr  
 450 455 460  
 Ser Leu Arg Asp Gly Met Asn Leu Val Ser Tyr Glu Phe Val Ala Cys  
 465 470 475 480  
 Gln Glu Ala Lys Lys Gly Val Leu Ile Leu Ser Glu Phe Ala Gly Ala  
 485 490 495  
 Ala Gln Ser Leu Gly Ala Gly Ala Ile Leu Val Asn Pro Trp Asn Ile  
 500 505 510  
 Thr Glu Val Ala Ala Ser Ile Gly Gln Ala Leu Asn Met Thr Ala Glu  
 515 520 525  
 Glu Arg Glu Lys Arg His Arg His Asn Phe His His Val Lys Thr His  
 530 535 540  
 Thr Ala Gln Glu Trp Ala Glu Thr Phe Val Ser Glu Leu Asn Asp Thr  
 545 550 555 560  
 Val Ile Glu Ala Gln Leu Arg Ile Ser Lys Val Pro Pro Glu Leu Pro  
 565 570 575  
 Gln His Asp Ala Ile Gln Arg Tyr Ser Lys Ser Asn Asn Arg Leu Leu  
 580 585 590  
 Ile Leu Gly Phe Asn Ala Thr Leu Thr Glu Pro Val Asp Asn Gln Gly  
 595 600 605  
 Arg Arg Gly Asp Gln Ile Lys Glu Met Asp Leu Asn Leu His Pro Glu  
 610 615 620  
 Leu Lys Gly Pro Leu Lys Ala Leu Cys Ser Asp Pro Ser Thr Thr Ile  
 625 630 635 640  
 Val Val Leu Ser Gly Ser Ser Arg Ser Val Leu Asp Lys Asn Phe Gly  
 645 650 655

Glu Tyr Asp Met Trp Leu Ala Ala Glu Asn Gly Met Phe Leu Arg Leu	660	665	670
Thr Asn Gly Glu Trp Met Thr Thr Met Pro Glu His Leu Asn Met Glu	675	680	685
Trp Val Asp Ser Val Lys His Val Phe Lys Tyr Phe Thr Glu Arg Thr	690	695	700
Pro Arg Ser His Phe Glu Thr Arg Asp Thr Ser Leu Ile Trp Asn Tyr	705	710	715
Lys Tyr Ala Asp Ile Glu Phe Gly Arg Leu Gln Ala Arg Asp Leu Leu	725	730	735
Gln His Leu Trp Thr Gly Pro Ile Ser Asn Ala Ser Val Asp Val Val	740	745	750
Gln Gly Ser Arg Ser Val Glu Val Arg Ala Val Gly Val Thr Lys Gly	755	760	765
Ala Ala Ile Asp Arg Ile Leu Gly Glu Ile Val His Ser Lys Ser Met	770	775	780
Thr Thr Pro Ile Asp Tyr Val Leu Cys Ile Gly His Phe Leu Gly Lys	785	790	795
Asp Glu Asp Val Tyr Thr Phe Phe Glu Pro Glu Leu Pro Ser Asp Met	805	810	815
Pro Ala Ile Ala Arg Ser Arg Pro Ser Ser Asp Ser Gly Ala Lys Ser	820	825	830
Ser Ser Gly Asp Arg Arg Pro Pro Ser Lys Ser Thr His Asn Asn Asn	835	840	845
Lys Ser Gly Ser Lys Ser Ser Ser Ser Ser Asn Ser Asn Asn Asn Asn	850	855	860
Lys Ser Ser Gln Arg Ser Leu Gln Ser Glu Arg Lys Ser Gly Ser Asn	865	870	875
His Ser Leu Gly Asn Ser Arg Arg Pro Ser Pro Glu Lys Ile Ser Trp	885	890	895
Asn Val Leu Asp Leu Lys Gly Glu Asn Tyr Phe Ser Cys Ala Val Gly	900	905	910
Arg Thr Arg Thr Asn Ala Arg Tyr Leu Leu Gly Ser Pro Asp Asp Val	915	920	925
Val Cys Phe Leu Glu Lys Leu Ala Asp Thr Thr Ser Ser Pro			

(2) INFORMATION FOR SEQ ID NO: 41:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 300 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Oryza sativa*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 41:

```
ATAAACTTCC TCGGACCAAA GAAGAGCATG TTGGTTGTGT CGGAGTTTAT TGGTTGCTCA    60
CCTTCACTGA GTGGAGCCAT TCGTGTTAAC CCGTGGAATA TCGAGGCAAC TGCAGAGGCA    120
CTGAATGAGG CCATCTCAAT GTCAGAGCGT AAAAGCAGCT GAGGCACGAA AAACATTACC    180
GTTATGTCAG CACCCATGAT GTTGCATATT GGTCTAAGAG CTTTGTACAG GACCTGGAGA    240
GGGCTTGCAA GGATCACTTT AGGAAACCAT GCTGGGGCAT TGGATTGGAT TTCGCTCAGG    300
```

(2) INFORMATION FOR SEQ ID NO: 42:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 627 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Selaginella lepidophylla*

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 4..627
- (D) OTHER INFORMATION: /partial

(ix) FEATURE:

- (A) NAME/KEY: unsure
- (B) LOCATION: 337..339
- (D) OTHER INFORMATION: /note = stopcodon

(ix) FEATURE:

- (A) NAME/KEY: unsure
- (B) LOCATION: 385..387
- (D) OTHER INFORMATION: /note = stopcodon

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 42:

ATT ATG TGG GTG CAT GAT TAC CAC CTC TGT CTG GTC CCT CAG ATG ATC	48
Met Trp Val His Asp Tyr His Leu Cys Leu Val Pro Gln Met Ile	
1 5 10 15	
CGC CAA AAG CTG CCA GAT GTG CAG ATT GGC TTC TTC CTC CAC ACC GCT	96
Arg Gln Lys Leu Pro Asp Val Gln Ile Gly Phe Phe Leu His Thr Ala	
20 25 30	
TTT CCC TCG TCA GAG GTC TTC CGC TGC TTG GCC GCA CGA AAG GAG CTG	144
Phe Pro Ser Ser Glu Val Phe Arg Cys Leu Ala Ala Arg Lys Glu Leu	
35 40 45	
CTG GAC GGC ATG CTT GGT GCC AAC TTG GTT GCT TTC CAG ACG CCA GAG	192
Leu Asp Gly Met Leu Gly Ala Asn Leu Val Ala Phe Gln Thr Pro Glu	
50 55 60	
TAT GCA CAC CAC TTC CTC CAG ACG TGC AGT CGC ATT TCT CTG CTG AAG	240
Tyr Ala His His Phe Leu Gln Thr Cys Ser Arg Ile Ser Leu Leu Lys	
65 70 75	
CAA CCG AGG AAG GCG TTC AGC TCG TTT CGT CAA TGT CTG GTC ATA ATG	288
Gln Pro Arg Lys Ala Phe Ser Ser Phe Arg Gln Cys Leu Val Ile Met	
80 85 90 95	
CAA GAA GCG CTA CGA GGG TCA AGA AGG TCA TCG TTG CGC GTG ACA AGC	336
Gln Glu Ala Leu Arg Gly Ser Arg Arg Ser Ser Leu Arg Val Thr Ser	
100 105 110	
TGA CAA CAT CGC GTG TAC GCG AGA AGC TTC TGT CGT ACG AGC TGT TCT	384
Xaa Gln His Arg Val Tyr Ala Arg Ser Phe Cys Arg Thr Ser Cys Ser	
115 120 125	
TGA ACA AGA ACC CAC AGT GGA GGG ACA AGG TCG TTC TCA TTC AGG TTG	432
Xaa Thr Arg Thr His Ser Gly Gly Thr Arg Ser Phe Ser Phe Arg Leu	
130 135 140	
CGA CCT CCA CGA CTG AGG ATT CTG AGC TTG CTG CGA CCG TAT CCG AAA	480
Arg Pro Pro Arg Leu Arg Ile Leu Ser Leu Leu Arg Pro Tyr Pro Lys	
145 150 155	
TTG TTA CAC GTA TTG ACG CTG TGC ACT CGA CGC TCA CAC ACA CCC ACT	528
Leu Leu His Val Leu Thr Leu Cys Thr Arg Arg Ser His Thr Pro Thr	
160 165 170 175	
CGT CTT CCT CAG GCA AGA CAT TGC GTT CTC GCA GTA CCT CGC ACT TCT	576
Arg Leu Pro Gln Ala Arg His Cys Val Leu Ala Val Pro Arg Thr Ser	
180 185 190	
CTC GAT CGC CGA TGC TCT TGC AAT CAA CTG TTC GAT GGC ATG AAC CTC	624
Leu Asp Arg Arg Cys Ser Cys Asn Gln Leu Phe Asp Gly Met Asn Leu	
195 200 205	
GTC	627
Val	

(2) INFORMATION FOR SEQ ID NO: 43:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 208 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 43:

Met	Trp	Val	His	Asp	Tyr	His	Leu	Cys	Leu	Val	Pro	Gln	Met	Ile	Arg	
1				5					10					15		
Gln	Lys	Leu	Pro	Asp	Val	Gln	Ile	Gly	Phe	Phe	Leu	His	Thr	Ala	Phe	
			20					25					30			
Pro	Ser	Ser	Glu	Val	Phe	Arg	Cys	Leu	Ala	Ala	Arg	Lys	Glu	Leu	Leu	
		35					40					45				
Asp	Gly	Met	Leu	Gly	Ala	Asn	Leu	Val	Ala	Phe	Gln	Thr	Pro	Glu	Tyr	
	50					55					60					
Ala	His	His	Phe	Leu	Gln	Thr	Cys	Ser	Arg	Ile	Ser	Leu	Leu	Lys	Gln	
65					70					75					80	
Pro	Arg	Lys	Ala	Phe	Ser	Ser	Phe	Arg	Gln	Cys	Leu	Val	Ile	Met	Gln	
				85					90					95		
Glu	Ala	Leu	Arg	Gly	Ser	Arg	Arg	Ser	Ser	Leu	Arg	Val	Thr	Ser	Xaa	
		100						105					110			
Gln	His	Arg	Val	Tyr	Ala	Arg	Ser	Phe	Cys	Arg	Thr	Ser	Cys	Ser	Xaa	
	115						120					125				
Thr	Arg	Thr	His	Ser	Gly	Gly	Thr	Arg	Ser	Phe	Ser	Phe	Arg	Leu	Arg	
	130					135						140				
Pro	Pro	Arg	Leu	Arg	Ile	Leu	Ser	Leu	Leu	Arg	Pro	Tyr	Pro	Lys	Leu	
145					150					155					160	
Leu	His	Val	Leu	Thr	Leu	Cys	Thr	Arg	Arg	Ser	His	Thr	Pro	Thr	Arg	
			165					170					175			
Leu	Pro	Gln	Ala	Arg	His	Cys	Val	Leu	Ala	Val	Pro	Arg	Thr	Ser	Leu	
		180						185					190			
Asp	Arg	Arg	Cys	Ser	Cys	Asn	Gln	Leu	Phe	Asp	Gly	Met	Asn	Leu	Val	
	195						200					205				

(2) INFORMATION FOR SEQ ID NO: 44:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 645 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Selaginella lepidophylla*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 44:

```
GGGTGGTTCT TGCACACGCC GTTTCCTCG TCTGAGATTT ACAGAACGCT GCCGCTGCGG 60
GCCGAGCTGC TCCAAGGCGT CTTAGGCGCG GACTTAGTGG GGTTCACAC ATACGACTAT 120
GCAAGGCACT TTGTTAGCGC GATGCACACG GATACTCGGG CTGGAAGGCA CTCCCAGGGT 180
GTCGAGGATC AAGGGAAGAT CACGCGAGTG GCTGCCTTCC CCGTGGATCG ATTCGGAGCG 240
ATTTATCGAC GCGTAGAGAC CGATGCGGTC AAGAAACACA TGCAAGAGCT GAGCCAGGTT 300
TTGCTGTCGT AAGGTTATGT TGGGGTGGAT AGGCTTGACA TGATTAAAGG AATTCCACAG 360
AAGCTGCTAG CCTTTGAAAA ATTCCTCGAG GAGAACTCCG AGTGGCGTGA TAAGGTCGTC 420
CTGGTGCAAA TCGCGGTGCC GACTAGAACG GACGTCCTCG AGTACCAAAA GCTTACGAGC 480
CAGGTTACAG AGATTGTTGG TCGCATAAAT GGACGTTTCG GCTCCTTGAC GGCTGTTCTT 540
ATCCATCACC TCGATCGGTC CATGAAATTT CCGGAGCTTT GTGCGTTATA TGCAATCACT 600
GATGTCCTGC TCGTGACATC CCTGCGCGAC GGCATGAACT TCGTC 645
```

(2) INFORMATION FOR SEQ ID NO: 45:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 498 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Arabidopsis thaliana*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 45:

```
GCCGTTGTGG ATTCATCGCC TCGACAAGC ACTCTTGTCTG TGTCTGAGTT TATTGGATGC 60
TCACCTTCTT TGAGTGGTGC CATTAGGGTG AATCCATGGG ATGTGGATGC TGTTGCTGAA 120
GCGGTAAACT CGGCTCTTAA AATAGTGAGA CTGAGAAGCA ACTACGGCAT GAGAAACATT 180
ATCATTATAT TAGCACTCAT GATGTTGGTT ATTGGGCAAA GAGCTTTATG CAGGATCTTG 240
```

AGAGAGCGTG CCGAGATCAT TATAGTAAAC GTTGTGTTGGG GATTGGTTTT GGCTTGGGGT 300  
 TCAGAGTTTT GTCACCTCTCT CCAAGTTTGA GGAAGCTATC TGTGGACACA TTTGTTCCAG 360  
 TTTATAGGAA AACCACAGAG AGGGCTAATA TTCTTTTATA ATGGTACTCT TTGTTCCGAA 420  
 AGCTCATTGT TCAAGATCCA GCAACGGGTT CCTTGTCTTA AGCCCCTTAA GGCCCCATAA 480  
 CCGGTGTTTT TTAGTGAG 498

(2) INFORMATION FOR SEQ ID NO: 46:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 463 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA to mRNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Arabidopsis thaliana
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 46:

GCCGTTGTGG ATTCATCGCC TCGACAAGC ACTCTTGTCG TGTCTGAGTT TATTGGATGC 60  
 TCACCTTCTT TGAGTGGTGC CATTGGGTGA ATCCATGGGA TGTGGATGCT GTTGCTGAAG 120  
 CGGTAAACTC GGCTCTTAAA ATGAGTGAGA CTGAGAAGCA ACTACGGCAT GAGAAACATT 180  
 ATCATTATAT TAGCACTCAT GATGTTGGTT ATTGGGCAAA GAGCTTTATG CAGGATCTTG 240  
 AGAGAGCGTG CCGAGATCAT TATAGTAAAC GTTGTGTTGGG GATTGGTTTT GGTGTTGGGGT 300  
 TCAGAGTTTT TGTCACCTCTC TCCAAGTTTA GGAAGCTATC TTGGGACAAT TGTTCCAGTT 360  
 TTTAGGGAAA ACACAGGGAA GGTTATTTCC TTGATTATAA TGGACCTTGT CCAAGCCCCA 420  
 TTTTAAAGGC CCAGGAACCG GGTTTTTTTT TCTTAAAGCC CCT 463

(2) INFORMATION FOR SEQ ID NO: 47:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 394 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA to mRNA
- (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Arabidopsis thaliana*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 47:

```
GGTATTGATG TAGAGGAAAT ACGTGGTGAA ATCGAAGAAA GCTGCAGGAG GATCAATGGA    60
GAGTTTGGGA AACCGGATAT CAACCTATCA TATATATTGA TACCCGGTTT CGATTAATGA    120
AATAAATGCT TATACCATAT TGCTGAGTGC GTGGTCGTTA CAGCTGTTAG AGATGGTATG    180
AACCTTACTC CCTACGAATA TATCGTTTGT AGACAAGGTT TACTTGGGTC TGAATCAGAC    240
TTTAGTGGCC CAAAGAAGAG CATGTTGGTT GCATCAAGTT TATTTGGATG TCCCCTTTCG    300
CTTAGTGGGG CTATACGCGT AAACCCATGG AACCGTTGAA GCTACTTGAG GAGCCTTAAT    360
TAGGCCCCCTC AAATATGCTG GAACACTACG GATG                                394
```

(2) INFORMATION FOR SEQ ID NO: 48:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 428 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Arabidopsis thaliana*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 48:

```
AAGTCCGTTG TGGATTCACG CCTCGCACAA GCACTCTTGT CGTGTCTAGT TTATTGGATG    60
CTCACCTTCT TTAGTGGTGC CATTAGGGTG AATCCATGGA TGTGGATGCT GTTGCTGAAG    120
CGGTAAACTC GGCTCTTAAA ATAGTGAGAC TGAGAAGCAA CTACGGCATG AGAAACATTA    180
TCATTATATT AGCACTCATG ATGTTGGTTA TTGGGCAAAG AGCTTTATGC AGGACTTAGA    240
GAGCGTGCCG AGATCATTAT AGTAAACGTT GTTGGGGGAT TGGTTTTGGT TTGGGGTTCA    300
AGTTTTGTCA CTCTCTCCAA GTTTTAGGAA GCTATCTTGT GGACACATTG TTCCAGTTTA    360
TAGAAACACA GGAAGGGGC TATATTCTTG TTAAATGGG ACCCCTTGTC CCTAAAAGTC    420
CCATTGTG                                428
```



(2) INFORMATION FOR SEQ ID NO: 49:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 481 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA to mRNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Arabidopsis thaliana*
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 49:

```
CAAACGAAGA GCTTCGTGGG AAAGTGGTTC TCGTGCAGAT TACTAATCCT GCTCGTAGTT   60
CAGGTAAGGA TGTTC AAGAT GTAGAGAAAC AGATAAATTT ATTGCTGATG AGATCAATTC  120
TAAATTTGGG AGACCTGGTG GTTATAAGCC TATTGTTTTG TAATGGACCT GTTAGTACTT  180
TGGATAAAGT TGCTTATTAC GCGATCTCGG AGTGTGTTGT CGTGAATCTG TGAGAGATGG  240
GATGAATTTG GTGCCTTATA AGTACACAGT GACTCGGCAA GGGAGCCCTG CTTTGGATGC  300
AGCTTTGGTT TTGGGGAGGA TGATGTTAGG AAGAGTGTGA TTATTGTTTC TGAGGTTCAA  360
CCGGTTGTCC TCCATCTCTA GTGGTGCGAT CCCTTTTAAT CCGTGGACAT CGATCAGCAC  420
TTACGCCATG AGCTTCAAAT CCGGTTTCCG CAAAGGGAAA ATTGCCCCGA GCTTAAGGCC  480
A                                                                                   481
```

(2) INFORMATION FOR SEQ ID NO: 50:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 395 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA to mRNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Arabidopsis thaliana*
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 50:

```
AGACCTGGTG GTTATAAGCC TATTGTGTTT GTCAATGGAC CTGTTAGTAC TTTGGATAAA   60
```

TTGCTTATTA CGCGATCTCG GAGTGTGTTG TCGTGAATCT GTGAGAGATG GGATGAATTT 120  
GGTGCCTTAT AAGTACACAG TGA CTCTCGGCA AGGGAGCCCT GCTTTGGATG CAGCTTTAGG 180  
TTTTGGGGAG GATGATGTTA GGAAGAGTGT GATTATTGTT TCTAGTTCAT CGGTTGTCTC 240  
CATCTCTGAG TGGTGCATC CGTTAATCCG TGAACATCG TGCAGTCACT AAACGCCATG 300  
AGCCTGCAAT ACGATGTCGC AAAGGGAAAA TCTTTGCCAC CAGAAGCATC ATAAGTACAT 360  
AAAGCCTCAC AATTGCCTAT TTGGGCCGGG GTTTT 395

(2) INFORMATION FOR SEQ ID NO: 51:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 431 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA to mRNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Oryza sativa*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION: 1
  - (D) OTHER INFORMATION: /standard\_name= "GENBANK ID: D22143"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 51:

GGGAATGGAG GGTCTCCGAG CTGCAGCAGC AATTTGAGGG GAAGACTGTG TTGCTCGGTG 60  
TGGATGACAT GGATATCTTC AAGGGTATCA ACTTGAAGCT TCTTGCCTTC GAGAATATGT 120  
TGAGGACACA TCCCAAGTGG CAGGGGCGGG CAGTGTGGT GCAAATTGCT AATCCGGCCC 180  
GTGGAAAGGG TAAGGATCTT GAAGCCATCC AGGCTGAGAT TCATGAGAGC TGCAAGAGGA 240  
TTAATGGAGA GTTTGGCCAG TCAGGATACA GCCCTGTTGT CTTCAATTGAC CGTGATGTGT 300  
CAAGTGTGGA GGAAGATTGC CTA CTACACA ATAGCAGAAT GTGTGGTGGT GACTGCTGTT 360  
AGGGATGGGA TTGACTTGAC ACCATATGGA TATATTGTCT GTAGGGCAGG GGTCTTACTC 420  
ACATCAGAGG T 431

(2) INFORMATION FOR SEQ ID NO: 52:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 496 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Oryza sativa*

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION: 1

(D) OTHER INFORMATION: /standard\_name= "GENBANK ID:  
D40048"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 52:

```
CTACCGTTCC CTCCCTGTTC GCGACGAGAT CCTCAAATCA CTGCTAAACT GCGATCTGAT    60
TGGGTTCCAC ACCTTTGATT ACGCGCGGCA TTTCCTGTCC TGCTGCAGCC GGATGCTGGG   120
GATCGAGTAC CAGTCGAAGA GGGGATATAT CGGTCTCGAT TACTTTGGCC GCACGTGTGG    180
GATAAAGATC ATGCCTGTTG GGATTAACAT GACGCAGCTG CAGACGCAGA TCCGGCTGCC    240
TGATCTTGAG TGGCGTGTCT CGAACTCCGG AAGCAGTTTG ATGGGAAGAC TGTCATGCTC    300
GGTGTGGATG ATATGGACAT ATTTAAGGGG ATTAATCTGA AAGTTCTTGC GTTTTGAGCA    360
GATGCTGAGG ACACACCCAA AATGGCAGCC AAGGCAGTTT TGGTGCAGAT TCAAACCAAG    420
GGTGGTTGTT GGGAGGACTT AGGTACAGCT AGATATGAGT TCAGGGGTAA TGACATTTCA    480
GGCGGTATTT CCTTGG                                                    496
```

(2) INFORMATION FOR SEQ ID NO: 53:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 288 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Oryza sativa*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 53:

GGACCAAAGA	AGAGCATGTT	GGTTGTGTCG	GAGTTTATTG	GTTGCTCACC	TTCCTGAGT	60
GGAGCCATTC	GTGTTAACCC	GTGGAATATC	GAGGCAACTG	CAGAGGCACT	GAATGAGGCC	120
ATCTCAATGT	CAGAGCGTAA	AAGCAGCTGA	GGCACGAAAA	ACATTACCGT	TATGTCAGCA	180
CCCATGATGT	TGCATATTGG	TCTAAGAGCT	TTGTACAGGA	CCTGGAGAGG	GCTTGCAAGG	240
ATCACTTTAG	GAAACCATGC	TGGGGCATTG	GATTGGATTT	CGCTCAGG		288

(2) INFORMATION FOR SEQ ID NO: 54:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2207 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Solanum tuberosum*  
(B) STRAIN: Kardal

(ix) FEATURE:

- (A) NAME/KEY: CDS  
(B) LOCATION: 161..1906

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature  
(B) LOCATION: 842..850  
(D) OTHER INFORMATION: /function= "putative glycosylationsite"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 54:

CTTTTCTGAG	TAATAACATA	GGCATTGATT	TTTTTTCAAT	TAATAACACC	TGCAAACATT	60										
CCCATTGCCG	GCATTCTCTG	TTCTTACAAA	AAAAAACATT	TTTTTGTTCA	CATAAATTAG	120										
TTATGGCATC	AGTATTGAAC	CCTTTAACTT	GTTATACAAT	ATG	GGT	AAA	GCT	ATA	175							
				Met	Gly	Lys	Ala	Ile								
				1				5								
ATT	TTT	ATG	ATT	TTT	ACT	ATG	TCT	ATG	AAT	ATG	ATT	AAA	GCT	GAA	ACT	223
Ile	Phe	Met	Ile	Phe	Thr	Met	Ser	Met	Asn	Met	Ile	Lys	Ala	Glu	Thr	
				10					15					20		
TGC	AAA	TCC	ATT	GAT	AAG	GGT	CCT	GTA	ATC	CCA	ACA	ACC	CCT	TTA	GTG	271
Cys	Lys	Ser	Ile	Asp	Lys	Gly	Pro	Val	Ile	Pro	Thr	Thr	Pro	Leu	Val	
				25				30					35			

ATT TTT CTT GAA AAA GTT CAA GAA GCT GCT CTT CAA ACT TAT GGC CAT	319
Ile Phe Leu Glu Lys Val Gln Glu Ala Ala Leu Gln Thr Tyr Gly His	
40 45 50	
AAA GGG TTT GAT GCT AAA CTG TTT GTT GAT ATG TCA CTG AGA GAG AGT	367
Lys Gly Phe Asp Ala Lys Leu Phe Val Asp Met Ser Leu Arg Glu Ser	
55 60 65	
CTT TCA GAA ACA GTT GAA GCT TTT AAT AAG CTT CCA AGA GTT GTG AAT	415
Leu Ser Glu Thr Val Glu Ala Phe Asn Lys Leu Pro Arg Val Val Asn	
70 75 80 85	
GGT TCA ATA TCA AAA AGT GAT TTG GAT GGT TTT ATA GGT AGT TAC TTG	463
Gly Ser Ile Ser Lys Ser Asp Leu Asp Gly Phe Ile Gly Ser Tyr Leu	
90 95 100	
AGT AGT CCT GAT AAG GAT TTG GTT TAT GTT GAG CCT ATG GAT TTT GTG	511
Ser Ser Pro Asp Lys Asp Leu Val Tyr Val Glu Pro Met Asp Phe Val	
105 110 115	
GCT GAG CCT GAA GGC TTT TTG CCA AAG GTG AAG AAT TCT GAG GTG AGG	559
Ala Glu Pro Glu Gly Phe Leu Pro Lys Val Lys Asn Ser Glu Val Arg	
120 125 130	
GCA TGG GCA TTG GAG GTG CAT TCA CTT TGG AAG AAT TTA AGT AGG AAA	607
Ala Trp Ala Leu Glu Val His Ser Leu Trp Lys Asn Leu Ser Arg Lys	
135 140 145	
GTG GCT GAT CAT GTA TTG GAA AAA CCA GAG TTG TAT ACT TTG CTT CCA	655
Val Ala Asp His Val Leu Glu Lys Pro Glu Leu Tyr Thr Leu Leu Pro	
150 155 160 165	
TTG AAA AAT CCA GTT ATT ATA CCG GGA TCG CGT TTT AAG GAG GTT TAT	703
Leu Lys Asn Pro Val Ile Ile Pro Gly Ser Arg Phe Lys Glu Val Tyr	
170 175 180	
TAT TGG GAT TCT TAT TGG GTA ATA AGG GGT TTG TTA GCA AGC AAA ATG	751
Tyr Trp Asp Ser Tyr Trp Val Ile Arg Gly Leu Leu Ala Ser Lys Met	
185 190 195	
TAT GAA ACT GCA AAA GGG ATT GTG ACT AAT CTG GTT TCT CTG ATA GAT	799
Tyr Glu Thr Ala Lys Gly Ile Val Thr Asn Leu Val Ser Leu Ile Asp	
200 205 210	
CAA TTT GGT TAT GTT CTT AAC GGT GCA AGA GCA TAC TAC AGT AAC AGA	847
Gln Phe Gly Tyr Val Leu Asn Gly Ala Arg Ala Tyr Tyr Ser Asn Arg	
215 220 225	
AGT CAG CCT CCT GTC CTG GCC ACG ATG ATT GTT GAC ATA TTC AAT CAG	895
Ser Gln Pro Pro Val Leu Ala Thr Met Ile Val Asp Ile Phe Asn Gln	
230 235 240 245	
ACA GGT GAT TTA AAT TTG GTT AGA AGA TCC CTT CCT GCT TTG CTC AAG	943
Thr Gly Asp Leu Asn Leu Val Arg Arg Ser Leu Pro Ala Leu Leu Lys	
250 255 260	

GAG AAT CAT TTT TGG AAT TCA GGA ATA CAT AAG GTG ACT ATT CAA GAT	991
Glu Asn His Phe Trp Asn Ser Gly Ile His Lys Val Thr Ile Gln Asp	
265 270 275	
GCT CAG GGA TCA AAC CAC AGC TTG AGT CGG TAC TAT GCT ATG TGG AAT	1039
Ala Gln Gly Ser Asn His Ser Leu Ser Arg Tyr Tyr Ala Met Trp Asn	
280 285 290	
AAG CCC CGT CCA GAA TCG TCA ACT ATA GAC AGT GAA ACA GCT TCC GTA	1087
Lys Pro Arg Pro Glu Ser Ser Thr Ile Asp Ser Glu Thr Ala Ser Val	
295 300 305	
CTC CCA AAT ATA TGT GAA AAA AGA GAA TTA TAC CGT GAA CTG GCA TCA	1135
Leu Pro Asn Ile Cys Glu Lys Arg Glu Leu Tyr Arg Glu Leu Ala Ser	
310 315 320 325	
GCT GCT GAA AGT GGA TGG GAT TTC AGT TCA AGA TGG ATG AGC AAC GGA	1183
Ala Ala Glu Ser Gly Trp Asp Phe Ser Ser Arg Trp Met Ser Asn Gly	
330 335 340	
TCT GAT CTG ACA ACA ACT AGT ACA ACA TCA ATT CTA CCA GTT GAT TTG	1231
Ser Asp Leu Thr Thr Thr Ser Thr Thr Ser Ile Leu Pro Val Asp Leu	
345 350 355	
AAT GCA TTC CTT CTG AAG ATG GAA CTT GAC ATT GCC TTT CTA GCA AAT	1279
Asn Ala Phe Leu Leu Lys Met Glu Leu Asp Ile Ala Phe Leu Ala Asn	
360 365 370	
CTT GTT GGA GAA AGT AGC ACG GCT TCA CAT TTT ACA GAA GCT GCT CAA	1327
Leu Val Gly Glu Ser Ser Thr Ala Ser His Phe Thr Glu Ala Ala Gln	
375 380 385	
AAT AGA CAG AAG GCT ATA AAC TGT ATC TTT TGG AAC GCA GAG ATG GGG	1375
Asn Arg Gln Lys Ala Ile Asn Cys Ile Phe Trp Asn Ala Glu Met Gly	
390 395 400 405	
CAA TGG CTT GAT TAC TGG CTT ACC AAC AGC GAC ACA TCT GAG GAT ATT	1423
Gln Trp Leu Asp Tyr Trp Leu Thr Asn Ser Asp Thr Ser Glu Asp Ile	
410 415 420	
TAT AAA TGG GAA GAT TTG CAC CAG AAC AAG AAG TCA TTT GCC TCT AAT	1471
Tyr Lys Trp Glu Asp Leu His Gln Asn Lys Lys Ser Phe Ala Ser Asn	
425 430 435	
TTT GTT CCG CTG TGG ACT GAA ATT TCT TGT TCA GAT AAT AAT ATC ACA	1519
Phe Val Pro Leu Trp Thr Glu Ile Ser Cys Ser Asp Asn Asn Ile Thr	
440 445 450	
ACT CAG AAA GTA GTT CAA AGT CTC ATG AGC TCG GGC TTG CTT CAG CCT	1567
Thr Gln Lys Val Val Gln Ser Leu Met Ser Ser Gly Leu Leu Gln Pro	
455 460 465	
GCA GGG ATT GCA ATG ACC TTG TCT AAT ACT GGA CAG CAA TGG GAT TTT	1615
Ala Gly Ile Ala Met Thr Leu Ser Asn Thr Gly Gln Gln Trp Asp Phe	
470 475 480 485	

CCG AAT GGT TGG CCC CCC CTT CAA CAC ATA ATC ATT GAA GGT CTC TTA	1663
Pro Asn Gly Trp Pro Pro Leu Gln His Ile Ile Ile Glu Gly Leu Leu	
490 495 500	
AGG TCT GGA CTA GAA GAG GCA AGA ACC TTA GCA AAA GAC ATT GCT ATT	1711
Arg Ser Gly Leu Glu Glu Ala Arg Thr Leu Ala Lys Asp Ile Ala Ile	
505 510 515	
CGC TGG TTA AGA ACT AAC TAT GTG ACT TAC AAG AAA ACC GGT GCT ATG	1759
Arg Trp Leu Arg Thr Asn Tyr Val Thr Tyr Lys Lys Thr Gly Ala Met	
520 525 530	
TAT GAA AAA TAT GAT GTC ACA AAA TGT GGA GCA TAT GGA GGT GGT GGT	1807
Tyr Glu Lys Tyr Asp Val Thr Lys Cys Gly Ala Tyr Gly Gly Gly Gly	
535 540 545	
GAA TAT ATG TCC CAA ACG GGT TTC GGA TGG TCA AAT GGC GTT GTA CTG	1855
Glu Tyr Met Ser Gln Thr Gly Phe Gly Trp Ser Asn Gly Val Val Leu	
550 555 560 565	
GCA CTT CTA GAG GAA TTT GGA TGG CCT GAA GAT TTG AAG ATT GAT TGC	1903
Ala Leu Leu Glu Glu Phe Gly Trp Pro Glu Asp Leu Lys Ile Asp Cys	
570 575 580	
TAATGAGCAA GTAGAAAAGC CAAATGAAAC ATCATTGAGT TTTATTTTCT TCTTTTGTTA	1963
AAATAAGCTG CAATGGTTTG CTGATAGTTT ATGTTTTTGTA TTACTATTTT ATAAGGTTTT	2023
TGTACCATAT CAAGTGATAT TACCATGAAC TATGTCGTTC GGACTCTTCA AATCGGATTT	2083
TGCAAAAATA ATGCAGTTTT GGAGAATCCG ATAACATAGA CCATGTATGG ATCTAAATTG	2143
TAAACAGCTT ACTATATTAA GTAAAAGAAA GATGATTCCT CTGCTTTAAA AAAAAAAAAA	2203
AAAA	2207

(2) INFORMATION FOR SEQ ID NO: 55:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 581 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 55:

Met Gly Lys Ala Ile Ile Phe Met Ile Phe Thr Met Ser Met Asn Met
1 5 10 15
Ile Lys Ala Glu Thr Cys Lys Ser Ile Asp Lys Gly Pro Val Ile Pro
20 25 30
Thr Thr Pro Leu Val Ile Phe Leu Glu Lys Val Gln Glu Ala Ala Leu
35 40 45

Gln	Thr	Tyr	Gly	His	Lys	Gly	Phe	Asp	Ala	Lys	Leu	Phe	Val	Asp	Met	50	55	60	
Ser	Leu	Arg	Glu	Ser	Leu	Ser	Glu	Thr	Val	Glu	Ala	Phe	Asn	Lys	Leu	65	70	75	80
Pro	Arg	Val	Val	Asn	Gly	Ser	Ile	Ser	Lys	Ser	Asp	Leu	Asp	Gly	Phe	85	90	95	
Ile	Gly	Ser	Tyr	Leu	Ser	Ser	Pro	Asp	Lys	Asp	Leu	Val	Tyr	Val	Glu	100	105	110	
Pro	Met	Asp	Phe	Val	Ala	Glu	Pro	Glu	Gly	Phe	Leu	Pro	Lys	Val	Lys	115	120	125	
Asn	Ser	Glu	Val	Arg	Ala	Trp	Ala	Leu	Glu	Val	His	Ser	Leu	Trp	Lys	130	135	140	
Asn	Leu	Ser	Arg	Lys	Val	Ala	Asp	His	Val	Leu	Glu	Lys	Pro	Glu	Leu	145	150	155	160
Tyr	Thr	Leu	Leu	Pro	Leu	Lys	Asn	Pro	Val	Ile	Ile	Pro	Gly	Ser	Arg	165	170	175	
Phe	Lys	Glu	Val	Tyr	Tyr	Trp	Asp	Ser	Tyr	Trp	Val	Ile	Arg	Gly	Leu	180	185	190	
Leu	Ala	Ser	Lys	Met	Tyr	Glu	Thr	Ala	Lys	Gly	Ile	Val	Thr	Asn	Leu	195	200	205	
Val	Ser	Leu	Ile	Asp	Gln	Phe	Gly	Tyr	Val	Leu	Asn	Gly	Ala	Arg	Ala	210	215	220	
Tyr	Tyr	Ser	Asn	Arg	Ser	Gln	Pro	Pro	Val	Leu	Ala	Thr	Met	Ile	Val	225	230	235	240
Asp	Ile	Phe	Asn	Gln	Thr	Gly	Asp	Leu	Asn	Leu	Val	Arg	Arg	Ser	Leu	245	250	255	
Pro	Ala	Leu	Leu	Lys	Glu	Asn	His	Phe	Trp	Asn	Ser	Gly	Ile	His	Lys	260	265	270	
Val	Thr	Ile	Gln	Asp	Ala	Gln	Gly	Ser	Asn	His	Ser	Leu	Ser	Arg	Tyr	275	280	285	
Tyr	Ala	Met	Trp	Asn	Lys	Pro	Arg	Pro	Glu	Ser	Ser	Thr	Ile	Asp	Ser	290	295	300	
Glu	Thr	Ala	Ser	Val	Leu	Pro	Asn	Ile	Cys	Glu	Lys	Arg	Glu	Leu	Tyr	305	310	315	320
Arg	Glu	Leu	Ala	Ser	Ala	Ala	Glu	Ser	Gly	Trp	Asp	Phe	Ser	Ser	Arg	325	330	335	
Trp	Met	Ser	Asn	Gly	Ser	Asp	Leu	Thr	Thr	Thr	Ser	Thr	Thr	Ser	Ile	340	345	350	



Leu Pro Val Asp Leu Asn Ala Phe Leu Leu Lys Met Glu Leu Asp Ile  
 355 360 365  
 Ala Phe Leu Ala Asn Leu Val Gly Glu Ser Ser Thr Ala Ser His Phe  
 370 375 380  
 Thr Glu Ala Ala Gln Asn Arg Gln Lys Ala Ile Asn Cys Ile Phe Trp  
 385 390 395 400  
 Asn Ala Glu Met Gly Gln Trp Leu Asp Tyr Trp Leu Thr Asn Ser Asp  
 405 410 415  
 Thr Ser Glu Asp Ile Tyr Lys Trp Glu Asp Leu His Gln Asn Lys Lys  
 420 425 430  
 Ser Phe Ala Ser Asn Phe Val Pro Leu Trp Thr Glu Ile Ser Cys Ser  
 435 440 445  
 Asp Asn Asn Ile Thr Thr Gln Lys Val Val Gln Ser Leu Met Ser Ser  
 450 455 460  
 Gly Leu Leu Gln Pro Ala Gly Ile Ala Met Thr Leu Ser Asn Thr Gly  
 465 470 475 480  
 Gln Gln Trp Asp Phe Pro Asn Gly Trp Pro Pro Leu Gln His Ile Ile  
 485 490 495  
 Ile Glu Gly Leu Leu Arg Ser Gly Leu Glu Glu Ala Arg Thr Leu Ala  
 500 505 510  
 Lys Asp Ile Ala Ile Arg Trp Leu Arg Thr Asn Tyr Val Thr Tyr Lys  
 515 520 525  
 Lys Thr Gly Ala Met Tyr Glu Lys Tyr Asp Val Thr Lys Cys Gly Ala  
 530 535 540  
 Tyr Gly Gly Gly Gly Glu Tyr Met Ser Gln Thr Gly Phe Gly Trp Ser  
 545 550 555 560  
 Asn Gly Val Val Leu Ala Leu Leu Glu Glu Phe Gly Trp Pro Glu Asp  
 565 570 575  
 Leu Lys Ile Asp Cys  
 580

(2) INFORMATION FOR SEQ ID NO: 56:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 56:

CTCAGATCTG GCCACAAA

18

(2) INFORMATION FOR SEQ ID NO: 57:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 57:

GTGCTCGTCT GCAGGTGC

18